







```

Db      2341 GGGTAAGAGAAAGACAGAGGCTGGGGTGGGGCCACCTGCTGTTTAAACAGGCACTTTC 2400
Qy      2401 TCCCTCTCGGGGCTTATTTTGTTCAGAACTAGACAGAGTGTGAACTCCTTGTGA 2460
Db      2401 TCCCTCTCGGGGCTTATTTTGTTCAGAACTAGACAGAGTGTGAACTCCTTGTGA 2460
Qy      2461 GAGAGGCTGGGAATCCTCTTTCAGAGACTTAAATCCATTTATCCCTGGAAATGTCGTC 2520
Db      2461 GAGAGGCTGGGAATCCTCTTTCAGAGACTTAAATCCATTTATCCCTGGAAATGTCGTC 2520
Qy      2521 TGGCCAGTGAAGAGGCTGGCTTTCAGAGCTCCCTGACCCCGCGCTGCCCGCCCTCGG 2580
Db      2521 TGGCCAGTGAAGAGGCTGGCTTTCAGAGCTCCCTGACCCCGCGCTGCCCGCCCTCGG 2580
Qy      2581 GGTAAATGAGGACTTCTGGCCCAAGAGGTTTGAAGCCATACGCTGAGAGTGGGTTA 2640
Db      2581 GGTAAATGAGGACTTCTGGCCCAAGAGGTTTGAAGCCATACGCTGAGAGTGGGTTA 2640
Qy      2641 GAATGTAAACAGCTTTTAACTTGGGATTTAAGAGCTTTTAAAGGTAAATATCTCTGAAA 2700
Db      2641 GAATGTAAACAGCTTTTAACTTGGGATTTAAGAGCTTTTAAAGGTAAATATCTCTGAAA 2700
Qy      2701 GAAAAATGACGTAAACCAAGCGGTGTAATGAAAGCTGTTTAAATTAAGAAAGCGTGG 2760
Db      2701 GAAAAATGACGTAAACCAAGCGGTGTAATGAAAGCTGTTTAAATTAAGAAAGCGTGG 2760
Qy      2761 GGCATGAATCTCACTTCTGCGCAATGAGTCAACATAGTATCTTAACTTAATAGAT 2820
Db      2761 GGCATGAATCTCACTTCTGCGCAATGAGTCAACATAGTATCTTAACTTAATAGAT 2820
Qy      2821 ACTAAATATATATTTCACTTCACTTGAAGTGTAGTCTTCCCGCCCGCCCACTTTT 2880
Db      2821 ACTAAATATATATTTCACTTCACTTGAAGTGTAGTCTTCCCGCCCGCCCACTTTT 2880
Qy      2881 TCTTTTGTGAGGCAAGTGTGATCACTGAGGCGCAGAGTTGAGACCAAGCTGCGCAACAT 2940
Db      2881 TCTTTTGTGAGGCAAGTGTGATCACTGAGGCGCAGAGTTGAGACCAAGCTGCGCAACAT 2940
Qy      2941 AGCGAAACCCGATCTCTAATAAATAAATAAATTTGGCGGCGATGTGGCGCATGCTT 3000
Db      2941 AGCGAAACCCGATCTCTAATAAATAAATAAATTTGGCGGCGATGTGGCGCATGCTT 3000
Qy      3001 GTGGTCCCGAGTACTCGGAGGTTGAGGAGGAGAGTGTGAAATGACAGAGTGGAGG 3060
Db      3001 GTGGTCCCGAGTACTCGGAGGTTGAGGAGGAGAGTGTGAAATGACAGAGTGGAGG 3060
Qy      3061 TTGCAATGAGCAAGATTGTGCACTGCACTCGAGCTGGGCAACAGAGCAAGACTTGTTC 3120
Db      3061 TTGCAATGAGCAAGATTGTGCACTGCACTCGAGCTGGGCAACAGAGCAAGACTTGTTC 3120
Qy      3121 TC 3122
Db      3121 TC 3122

```

```

RESULT 2
AK124499 3122 bp mRNA linear PRI 09-SEP-2003
DEFINITION Homo sapiens cDNA FLJ42508 flB, clone BRACE2043665.
ACCESSION AK124499
VERSION AK124499.1 GI:34530295
KEYWORDS oligo cloning; flB (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

```

```

REFERENCE
AUTHORS
1 Kawakami, B., Sugiyama, A., Takemoto, M., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
Matsumoto, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,

```

```

TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Oshima, A., Suzuki, Y., Sugano, S., Nagahara, K., Masuno, Y., Nagai, K.
and Isogai, T.
NEO human cDNA sequencing project
2 (bases 1 to 3122)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (15-JUN-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail: genomics@kazusa.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
Location/Qualifiers
1..3122
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRACE2043665"
/tissue_type="cerebellum"
/clone_id="BRACE2"
/feature="cloning vector: pME18P13"
858..1565
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC85865.1"
/db_xref="GI:34530295"
/translation="MARBECKALDELAKTACVHYLYTVGSSADSONLROELQTR
QKQBLASTCARLTVLRDGLADEPRLVNAASGCLDLBARDRLBEGGA
FPLAFLRPLVTVAGASGVAAALASLRLRLAEEDPVDADLRLERLVLQVEM
IDNMENKVNPRVTVQAOAAGABLLSTVASGPSVSLIOERGCGCDPRKLAAILFG
AVLAAVALAVCAVLGS"

```

## ORIGIN

```

Query Match 100.0%; Score 3122; DB 8; Length 3122;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      1 ACTAGAGGTGGGGGTTAGGCTTGAAGACCGACCAAGTGAAGGCAACGCGGACGAGC 60
Db      1 ACTAGAGGTGGGGGTTAGGCTTGAAGACCGACCAAGTGAAGGCAACGCGGACGAGC 60
Qy      61 ACTGACCCCGCGCGCGCGCCAGCCCTCGATGTCAGTCACTGCTGCTTGGGAGCAG 120
Db      61 ACTGACCCCGCGCGCGCGCCAGCCCTCGATGTCAGTCACTGCTGCTTGGGAGCAG 120
Qy      121 GAGGTGCCAGTCTGCGGGGACCCGACGTCCTGTCCGCGACAGGCTCCGGAATCACT 180
Db      121 GAGGTGCCAGTCTGCGGGGACCCGACGTCCTGTCCGCGACAGGCTCCGGAATCACT 180
Qy      121 GAGGTGCCAGTCTGCGGGGACCCGACGTCCTGTCCGCGACAGGCTCCGGAATCACT 180
Db      121 GAGGTGCCAGTCTGCGGGGACCCGACGTCCTGTCCGCGACAGGCTCCGGAATCACT 180
Qy      161 ATAGCTGAGTCTAGTCCCATCAAGGAAATCTCGCGGAGGCTGCGCGCTTTTAA 240
Db      161 ATAGCTGAGTCTAGTCCCATCAAGGAAATCTCGCGGAGGCTGCGCGCTTTTAA 240
Qy      241 CTTGGGCTCTAGTTTCCCATCCGTAATAGAAAGGCTTGGATCTCCGAGCGCTAAACA 300
Db      241 CTTGGGCTCTAGTTTCCCATCCGTAATAGAAAGGCTTGGATCTCCGAGCGCTAAACA 300
Qy      241 CTTGGGCTCTAGTTTCCCATCCGTAATAGAAAGGCTTGGATCTCCGAGCGCTAAACA 300
Db      241 CTTGGGCTCTAGTTTCCCATCCGTAATAGAAAGGCTTGGATCTCCGAGCGCTAAACA 300
Qy      301 TTCCGAATCTGAGTGGGCGGAGGAGGAGGAGGATGGGCAACCAACGTAACCTCCC 360
Db      301 TTCCGAATCTGAGTGGGCGGAGGAGGAGGAGGAGGATGGGCAACCAACGTAACCTCCC 360
Qy      361 CGCGTGAAGCCCGGCTCACTGATCCAGGGGGTGGAGCTCCGCGGAGGAGGAGGAG 420
Db      361 CGCGTGAAGCCCGGCTCACTGATCCAGGGGGTGGAGCTCCGCGGAGGAGGAGGAG 420
Qy      421 GTGGGCGGCTCTAGAAACCTTACCGGCGCGCTTGGCAGCGCTTAAAGCGAGCGCG 480
Db      421 GTGGGCGGCTCTAGAAACCTTACCGGCGCGCTTGGCAGCGCTTAAAGCGAGCGCG 480

```

Db	421	GTGGGCGGGTCTTAGAAAACTTACCCGCGCCTTGTGACGCGCTTAAGCGGAGCGC	480
QY	481	CGGCTCTGACGCTGCTTGGCCCGGAGTTTGACCTCAAGAGATGTGGACCGGACCTTC	540
Db	481	CGGCTCTGACGCTGCTTGGCCCGGAGTTTGACCTCAAGAGATGTGGACCGGACCTTC	540
QY	541	AGCTTCGAGGGAGGACCCGTGTGAAGGCCAGGGCGGTGTGAGAGACACACGTGTGACTTCG	600
Db	541	AGCTTCGAGGGAGGACCCGTGTGAAGGCCAGGGCGGTGTGAGAGACACACGTGTGACTTCG	600
QY	601	AGTGGCCTTGGGAGAGATGGAACAGAGGAGCGGGGACCGCTTAAGGGGCTCCCTTGCGC	660
Db	601	AGTGGCCTTGGGAGAGATGGAACAGAGGAGCGGGGACCGCTTAAGGGGCTCCCTTGCGC	660
QY	661	GCCCGGTCCGAGAGGCGCACGTTCAGGGGTCCCGGCGGGCTCCGTGACGTTTGACGTGA	720
Db	661	GCCCGGTCCGAGAGGCGCACGTTCAGGGGTCCCGGCGGGGTCCGTGACGTTTGACGTGA	720
QY	721	GCGCCGAGGAGTCACGGACCATGAAAGCCTTGCTGCGCGCGGCCCAAGCCGGATG	780
Db	721	GCGCCGAGGAGTCACGGACCATGAAAGCCTTGCTGCGCGCGGCCCAAGCCGGATG	780
QY	781	GGGGTTAGCCCATCTCTGCGCGGCTGAAGGGGAGGCTTAAAGGGGCGGGGCGGGCGCC	840
Db	781	GGGGTTAGCCCATCTCTGCGCGGCTGAAGGGGAGGCTTAAAGGGGCGGGGCGGGCGCC	840
QY	841	AGCGGAGACCCACCGCGATGTGCGAGAGGAGAGTGCAGAGCGCTGTGACGGGCTCAACA	900
Db	841	AGCGGAGACCCACCGCGATGTGCGAGAGGAGAGTGCAGAGCGCTGTGACGGGCTCAACA	900
QY	901	AGACGACTGCGTGTCAACCACTTGTGTGACCTGTGTGCTCGGCGGACTTCGACAGA	960
Db	901	AGACGACTGCGTGTCAACCACTTGTGTGACCTGTGTGCTCGGCGGACTTCGACAGA	960
QY	961	ACCTGGCGGAGAGAGCTGTGAAAAGAGCGGCGAGAGGCGGAGAGTGGCGGTGCACCT	1020
Db	961	ACCTGGCGGAGAGAGCTGTGAAAAGAGCGGCGAGAGGCGGAGAGTGGCGGTGCACCT	1020
QY	1021	GCGCCCGGCTGACTGTGTGCTGCGCGACCGGAGGCTTGAGCGCGACGAGCGGCGAGT	1080
Db	1021	GCGCCCGGCTGACTGTGTGCTGCGCGACCGGAGGCTTGAGCGCGACGAGCGGCGAGT	1080
QY	1081	TCGAGCGGCTCTGGGTGGCTTCTCGGCTGTGCTGAACTTGCTGAGCGGACATGCGAC	1140
Db	1081	TCGAGCGGCTCTGGGTGGCTTCTCGGCTGTGCTGAACTTGCTGAGCGGACATGCGAC	1140
QY	1141	GCTGCGTGAAGCTGTGGCGCGCGGTCCCGCTGACAGCGCGCGGAGACCGCTGTGGCGCA	1200
Db	1141	GCTGCGTGAAGCTGTGGCGCGCGGTCCCGCTGACAGCGCGCGGAGACCGCTGTGGCGCA	1200
QY	1201	CAGGTGTGAGCTGACGCTCTTCGCGGCTGTGCGGCGCGGCGCTGAGACACCGGACGCTTCG	1260
Db	1201	CAGGTGTGAGCTGACGCTCTTCGCGGCTGTGCGGCGCGGCGCTGAGACACCGGACGCTTCG	1260
QY	1261	GCGTTCGAGGCGGAGGAGCGACTTTCAGCGTTCGCGGACTTCGCGGAGCTTGAGCGCGAGTCC	1320
Db	1261	GCGTTCGAGGCGGAGGAGCGACTTTCAGCGTTCGCGGACTTCGCGGAGCTTGAGCGCGAGTCC	1320
QY	1321	TTTCAAGTGTGGGCGAGATGATTCGACACATGTGAAGATTAAGTCAACTGCGCCCGCTGACACCG	1380
Db	1321	TTTCAAGTGTGGGCGAGATGATTCGACACATGTGAAGATTAAGTCAACTGCGCCCGCTGACACCG	1380
QY	1381	TGCAAGCCCGGAGGCGGCGGAGGCTCTGTCCAGGATCAAGGATCAAGCGCGGCGCTCTCT	1440
Db	1381	TGCAAGCCCGGAGGCGGCGGAGGCTCTGTCCAGGATCAAGGATCAAGCGCGGCGCTCTCT	1440
QY	1441	CGGTGTGTCTTTGACGAGAGCGCGGGGGGGGTGTGACCCCAAGAAAGCCCTGTGGCGCA	1500
Db	1441	CGGTGTGTCTTTGACGAGAGCGCGGGGGGGGTGTGACCCCAAGAAAGCCCTGTGGCGCA	1500
QY	1501	TCCTTTTGGGCGCGTGTCTGTGGCGGCTGTGGCGCTTGAAGCGGTGTGCGGAGGCTGA	1560
Db	1501	TCCTTTTGGGCGCGTGTCTGTGGCGGCTGTGGCGCTTGAAGCGGTGTGCGGAGGCTGA	1560

QY	1561	CTGTACAGACACCCGAGCGCCGCTGCTGCTGCCCTCCCTCCGCTGAGAAAGAAACATCCGG	1620
Db	1561	GCTACAGACACCCGAGCGCCGCTGCTGCTGCCCTCCCTCCCTGAGAAAGAAACATCCGG	1620
QY	1621	GATGGGTGTGGGGGTCTGGCTGTGCAAGGGGAGTGATCTTAAACCCTGTGTGTGATGG	1680
Db	1621	GATGGGTGTGGGGGTCTGGCTGTGCAAGGGGAGTGATCTTAAACCCTGTGTGTGATGG	1680
QY	1681	GTACACCGCGCTTCCAGTGCACATCTGCTGGGCAAGACACGGTTTTCTCTTGGCTGGC	1740
Db	1681	GTACACCGCGCTTCCAGTGCACATCTGCTGGGCAAGACACGGTTTTCTCTTGGCTGGC	1740
QY	1741	CCGGGAGAAAGTTAACTTTGGCGCCGCGCGCTGACGGGCAATTACCGCTAAAGTCTGACAGAGCT	1800
Db	1741	CCGGGAGAAAGTTAACTTTGGCGCCGCGCGCTGACGGGCAATTACCGCTAAAGTCTGACAGAGCT	1800
QY	1801	TTATTTCCCTATTAAATGAAAAACGGTGACAGTGACCTTAAGATCCCTCCGAGTTAATGATT	1860
Db	1801	TTATTTCCCTATTAAATGAAAAACGGTGACAGTGACCTTAAGATCCCTCCGAGTTAATGATT	1860
QY	1861	AAACAATGTGTGTGTGGGGCGTCTTTACAGGAGATCCGAGTGGGTGCCACCCCTGACCA	1920
Db	1861	AAACAATGTGTGTGTGGGGCGTCTTTACAGGAGATCCGAGTGGGTGCCACCCCTGACCA	1920
QY	1921	GCGTCCGCCCCCTTCTGTCGTTGGGACAGTTTGAAGATGGGTGGGTGTGAGTGAAGTTTG	1980
Db	1921	GCGTCCGCCCCCTTCTGTCGTTGGGACAGTTTGAAGATGGGTGGGTGTGAGTGAAGTTTG	1980
QY	1981	GAGAGGGACGCTGTTGGTCTATGTGTGTGTCTGTGTTCCCGGACAAAGAAAAATTCGAA	2040
Db	1981	GAGAGGGACGCTGTTGGTCTATGTGTGTGTCTGTGTTCCCGGACAAAGAAAAATTCGAA	2040
QY	2041	TCAATATGACAGACCTTTATTTACCTTAATCTTTCAGGGCCCTAAATTATGAGAGTGTCC	2100
Db	2041	TCAATATGACAGACCTTTATTTACCTTAATCTTTCAGGGCCCTAAATTATGAGAGTGTCC	2100
QY	2101	TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCTCTAGCAGAGT	2160
Db	2101	TGAGAGCAGTTCATACAAAGGGCTTTCTCTAAGACGCGCTACAGCCCTTCTCTAGCAGAGT	2160
QY	2161	TTATTCOAATCGTCCCAAGAGCAGCTAGAAAGATTGAAGGTCAATGACCTCCACATGCGG	2220
Db	2161	TTATTCOAATCGTCCCAAGAGCAGCTAGAAAGATTGAAGGTCAATGACCTCCACATGCGG	2220
QY	2221	CTCAGAGGGCTGACCCCTATTTAGAAGAACCAAGAGGTGGTGAACCTACTCTCACGGAC	2280
Db	2221	CTCAGAGGGCTGACCCCTATTTAGAAGAACCAAGAGGTGGTGAACCTACTCTCACGGAC	2280
QY	2281	TTTGATTCAGTGTGCGCACATTTGGCTGTGGGAAAAAGGCTCTTCCCAAGCACCCCGAAGATGG	2340
Db	2281	TTTGATTCAGTGTGCGCACATTTGGCTGTGGGAAAAAGGCTCTTCCCAAGCACCCCGAAGATGG	2340
QY	2341	GGGTAAAGAAAGACAGAGGCTTGGGGTAGGGGCACCTGGTGTTTAAACAGGACACTTTC	2400
Db	2341	GGGTAAAGAAAGACAGAGGCTTGGGGTAGGGGCACCTGGTGTTTAAACAGGACACTTTC	2400
QY	2401	TCCTTCTCTGGGGCTTAATTTTGTTCAGAACTAGAACCAAGATGTTTGAACCTCCTTTGCA	2460
Db	2401	TCCTTCTCTGGGGCTTAATTTTGTTCAGAACTAGAACCAAGATGTTTGAACCTCCTTTGCA	2460
QY	2461	GGAAGGCTGGGAAATCCCTCTTAAAGGACTTAATCCCTAATTATGCCCTGGAAATGTGGCGTGC	2520
Db	2461	GGAAGGCTGGGAAATCCCTCTTAAAGGACTTAATCCCTAATTATGCCCTGGAAATGTGGCGTGC	2520
QY	2521	TGGCCAGTAGAGAGGGCTGGCTTTGGCAGCTCCCTGACCCCGCGCTGCGCCGCTCCGCG	2580
Db	2521	TGGCCAGTAGAGAGGGCTGGCTTTGGCAGCTCCCTGACCCCGCGCTGCGCCGCTCCGCG	2580
QY	2581	GGTAATGTGGCAATTACTGGGCCACAGAGTTTTTGAAGCAATCAAGCTCTGAGACTGGGGTTA	2640
Db	2581	GGTAATGTGGCAATTACTGGGCCACAGAGTTTTTGAAGCAATCAAGCTCTGAGACTGGGGTTA	2640

QY 2641 GAATGATACAGCTTTAACTTGGATTGAAGAGCTTTTAAAGATATATCTCTGAAA 2700  
Db 2641 GAATGATACAGCTTTAACTTGGATTGAAGAGCTTTTAAAGATATATCTCTGAAA 2700  
QY 2701 GAAAAATGACGTACCAAGCGGTGATCTATGAAAGCTGTTATTTATTAAGAAGCTGG 2760  
Db 2701 GAAAAATGACGTACCAAGCGGTGATCTATGAAAGCTGTTATTTATTAAGAAGCTGG 2760  
QY 2761 GCCATGACCTACCTGCGCAATGAGTCAAAACATGATCTTTATGATCTTATGATT 2820  
Db 2761 GCCATGACCTACCTGCGCAATGAGTCAAAACATGATCTTTATGATCTTATGATT 2820  
QY 2821 ACTAATATATATTTTATCTACTTCTGAAGTTGATGTCCTCCCGCCCGCACTTTT 2880  
Db 2821 ACTAATATATATTTTATCTACTTCTGAAGTTGATGTCCTCCCGCCCGCACTTTT 2880  
QY 2881 TCTTTTGAAGGAGGTGATCACTGAGGCCAGAGTTTGAGACCAAGCCTGGCAACAT 2940  
Db 2881 TCTTTTGAAGGAGGTGATCACTGAGGCCAGAGTTTGAGACCAAGCCTGGCAACAT 2940  
QY 2941 AGCGAAACCCGATCTCTACTTAAATATATATTAATTTGGCCGCGCATGTCGCGCT 3000  
Db 2941 AGCGAAACCCGATCTCTACTTAAATATATTAATTTGGCCGCGCATGTCGCGCT 3000  
QY 3001 GTGCTCCAGCTACTCGGAGGTTGAGGAGAGAGTCCGTTGAATGACAGAGTGAAG 3060  
Db 3001 GTGCTCCAGCTACTCGGAGGTTGAGGAGAGAGTCCGTTGAATGACAGAGTGAAG 3060  
QY 3061 TTGCAATGAGCAAGATTGTGCCATGCTGCACTCCAGCTGGGCAACAGAGCAAGCTGTC 3120  
Db 3061 TTGCAATGAGCAAGATTGTGCCATGCTGCACTCCAGCTGGGCAACAGAGCAAGCTGTC 3120  
QY 3121 TC 3122  
Db 3121 TC 3122

RESULT 3  
AC008736/c 191925 bp DNA 1linear PRI 27-SEP-2000  
LOCUS AC008736 Homo sapiens chromosome 19 clone CTD-2538C1, complete sequence.  
DEFINITION AC008736  
VERSION AC008736.6 GI:10312244  
KEYWORDS HTG.

SOURCE  
ORGANISM Homo sapiens (human)

REFERENCE  
AUTHORS Homo sapiens  
TITLE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

REFERENCE  
AUTHORS DOB Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 191925)  
TITLE DOB Joint Genome Institute.

REFERENCE  
AUTHORS Direct Submission  
JOURNAL DOB Joint Genome Institute.

REFERENCE  
AUTHORS Submtted (03-AUG-1999) Production Sequencing Facility, DOB Joint  
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
TITLE 3 (bases 1 to 191925)

REFERENCE  
AUTHORS DOB Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submtted (27-SEP-2000) DOB Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA

REFERENCE  
AUTHORS On Sep 27, 2000 this sequence version replaced gi:8575905.  
JOURNAL Draft Sequence Produced by DOB Joint Genome Institute

REFERENCE  
AUTHORS www.jgi.doe.gov  
JOURNAL Finishing Completed at Stanford Human Genome Center

REFERENCE  
AUTHORS Quality: Phrap Quality >=40 99.9% of Sequence;  
JOURNAL Estimated Total Number of Errors is 0.1.

REFERENCE  
AUTHORS SHGC-57769 G37408,  
JOURNAL Location/Qualifiers

FEATURES

source 1..191925  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clone="CTD-2538C1"

Query Match 92.7%; Score 2892.8; DB 8; Length 191925;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2894; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTAAGAGTGGGGTTAGCGCTTGGAGAGCAACCAACCTGAGCGCAACGCGGACAGGAC 60  
Db 129460 ACTAAGAGTGGGGTTAGCGCTTGGAGAGCAACCAACCTGAGCGCAACGCGGACAGGAC 129401  
QY 61 ACCTGAACCCCGGCGGCGCCAGCCCTCGGATTTGCCAGTCACTGCTCGCTTTGGGCAAG 120  
Db 129400 ACCTGAACCCCGGCGGCGCCAGCCCTCGGATTTGCCAGTCACTGCTCGCTTTGGGCAAG 129341  
QY 121 GAGGTGCCAGTCTCGCGGGGAGACCCGAGCTGTCGCGCGGAGAGGAGTCCGGAGTCAAG 180  
Db 129340 GAGGTGCCAGTCTCGCGGGGAGACCCGAGCTGTCGCGCGGAGAGGAGTCCGGAGTCAAG 129281  
QY 181 ATAGCTGGGTTCTAGTCCATCAAGAGCAAAATCTCCGCGGAGAGCTGAGCCGCTTTTAA 240  
Db 129280 ATAGCTGGGTTCTAGTCCATCAAGAGCAAAATCTCCGCGGAGAGCTGAGCCGCTTTTAA 129221  
QY 241 CCGTGGGCTCAAGTTTCCCATCCGTAATGAAACGGGTTGATCTCCGAGCGCTAAACA 300  
Db 129220 CCGTGGGCTCAAGTTTCCCATCCGTAATGAAACGGGTTGATCTCCGAGCGCTAAACA 129161  
QY 301 TTCCAGAACTCGGATGAGGCGGAGAGGAGGAGTGGGCCCAACCTCAACCTGACCTCC 360  
Db 129160 TTCCAGAACTCGGATGAGGCGGAGAGGAGGAGTGGGCCCAACCTCAACCTGACCTCC 129101  
QY 361 CGCGTGAAGCCCGGCTACCACTGATCCAGGGGTTGACAGTCCGCGCGGAGAGCGAG 420  
Db 129100 CGCGTGAAGCCCGGCTACCACTGATCCAGGGGTTGACAGTCCGCGCGGAGAGCGAG 129041  
QY 421 GTGGCGGAGTCTTGAAGAAACCTTACCCGCGCGCTTGGCAGCGCTTAAGCGGAGCGG 480  
Db 129040 GTGGCGGAGTCTTGAAGAAACCTTACCCGCGCGCTTGGCAGCGCTTAAGCGGAGCGG 128981  
QY 481 CGGCTCTGAGCGCTCTTGGCCCGGAGTTGGACCAACGAGAGTGGGAGCGGAGCGG 540  
Db 128980 CGGCTCTGAGCGCTCTTGGCCCGGAGTTGGACCAACGAGAGTGGGAGCGGAGCGG 128921  
QY 541 AGCTTCGACGAGGAGCAACGTTGAGAGGCGGCGGTCAGAGACAGACGTTGACTCGG 600  
Db 128920 AGCTTCGACGAGGAGCAACGTTGAGAGGCGGCGGTCAGAGACAGACGTTGACTCGG 128861  
QY 601 AGTGGCGCTTGGGAGAGTGAAGAGAGGAGGAGGAGCGCTTAAAGGAGCTCCCTCTGCGC 660  
Db 128860 AGTGGCGCTTGGGAGAGTGAAGAGAGGAGGAGGAGCGCTTAAAGGAGCTCCCTCTGCGC 128801  
QY 661 GCCCGCTCGGAGAGGCGGAGCGTCAAGGAGTCCCGGCGGAGCTCCGTTGACGTTGGCGGTA 720  
Db 128800 GCCCGCTCGGAGAGGCGGAGCGTCAAGGAGTCCCGGCGGAGCTCCGTTGACGTTGGCGGTA 128741  
QY 721 GCGCGGAGAGTCAAGGAGCAATGAAGAGCGTTGCGCGCGGCGGCGGAGCGGAGATG 780  
Db 128740 GCGCGGAGAGTCAAGGAGCAATGAAGAGCGTTGCGCGCGGCGGCGGAGCGGAGATG 128681  
QY 781 GGGGTTAAGCAATCTGCGCGGCTGAGGAGGAGCTTAAAGGAGCGGCGGCGGCGGCGG 840  
Db 128680 GGGGTTAAGCAATCTGCGCGGCTGAGGAGGAGGAGCTTAAAGGAGCGGCGGCGGCGGCGG 128621  
QY 841 AGCGGAGCGGAGCGGAGTGGAGAGGAGGAGTGAAGGCGGCTGAGCGGCGGCTCAACA 900  
Db 128620 AGCGGAGCGGAGCGGAGTGGAGAGGAGGAGTGAAGGCGGCTGAGCGGCGGCTCAACA 128561  
QY 901 AGACGACTGCGTGTACCAACCACTGATGCTGAGCCGTCGCTGCGGCGGAGCTGCAAGA 960

Db	128560	AGA	GAC	CTG	CGT	GC	TAC	CA	CA	CC	TGG	TGT	GT	CA	CG	TC	GG	TGG	CT	CG	CG	GA	CT	CC	GA	128501	
Qy	961	ACT	GCG	CA	GGA	AG	CT	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	1020	
Db	128500	ACT	GCG	CA	GGA	AG	CT	G	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	A	128441	
Qy	1021	GG	CG	CG	CG	CT	GA	CT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	1080	
Db	128440	GG	CG	CG	CG	CT	GA	CT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	128381	
Qy	1081	TC	GA	GC	GC	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	1140	
Db	128380	TC	GA	GC	GC	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	CT	128321	
Qy	1141	GCT	GC	CT	GA	GA	CT	GG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1200	
Db	128320	GCT	GC	CT	GA	GA	CT	GG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	128261	
Qy	1201	CAG	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	1260	
Db	128260	CAG	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	GT	128201	
Qy	1261	GG	CT	CG	GA	GG	CG	GA	CT	CG	CG	GA	CT	CG	CG	GA	CT	CG	CG	GA	CT	CG	CG	GA	CT	CG	1320
Db	128200	GG	CT	CG	GA	GG	CG	GA	CT	CG	CG	GA	CT	CG	CG	GA	CT	CG	CG	GA	CT	CG	CG	GA	CT	CG	128141
Qy	1321	TT	CA	AG	GT	GG	CG	GA	TA	TC	GA	CA	CA	AT	GA	AG	TA	GA	AG	TA	GA	AG	TA	GA	AG	TA	1380
Db	128140	TT	CA	AG	GT	GG	CG	GA	TA	TC	GA	CA	CA	AT	GA	AG	TA	GA	AG	TA	GA	AG	TA	GA	AG	TA	128081
Qy	1381	TG	CA	AG	CC	CG	CA	AG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1440	
Db	128080	TG	CA	AG	CC	CG	CA	AG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	128021	
Qy	1441	CG	GT	CG	GT	CT	CT	CA	GA	AG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1500	
Db	128020	CG	GT	CG	GT	CT	CT	CA	GA	AG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	127961	
Qy	1501	TC	CT	TT	TC	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1560	
Db	127960	TC	CT	TT	TC	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	127901	
Qy	1561	GCT	GA	CA	GA	CA	CC	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1620	
Db	127900	GCT	GA	CA	GA	CA	CC	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	127841	
Qy	1621	GAT	GG	GT	GG	GT	GG	GT	GG	GT	GG	GT	GG	GT	GG	GT	GG	GT	GG	GT	GG	GT	GG	GT	GG	1680	
Db	127840	GAT	GG	GT	GG	GT	GG	GT	GG	GT	GG	GT	GG	GT	GG	GT	GG	GT	GG	GT	GG	GT	GG	GT	GG	127781	
Qy	1661	GT	AC	AC	GC	GC	GT	TT	TC	CA	GT	CA	GT	CT	CG	GA	CA	CG	GT	TT	CT	CT	TT	CT	CT	1740	
Db	127780	GT	AC	AC	GC	GC	GT	TT	TC	CA	GT	CA	GT	CT	CG	GA	CA	CG	GT	TT	CT	CT	TT	CT	CT	127721	
Qy	1741	CC	GG	GA	GA	GT	TA	CT	TT	GG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	1800	
Db	127720	CC	GG	GA	GA	GT	TA	CT	TT	GG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	CG	127661	
Qy	1801	TT	AT	T	CC	CT	TA	TA	TA	GA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	1860	
Db	127660	TT	AT	T	CC	CT	TA	TA	TA	GA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	127601	
Qy	1861	AA	CA	CA	TA	TA	TA	TA	TA	GA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	1920	
Db	127600	AA	CA	CA	TA	TA	TA	TA	TA	GA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	AA	127541	
Qy	1921	GC	GT	CG	CC	CC	CT	TT	CT	CG	GT	GA	CA	GT	TT	GA	AA	AG	TT	GG	GT	GG	GT	GG	GT	1980	
Db	127540	GC	GT	CG	CC	CC	CT	TT	CT	CG	GT	GA	CA	GT	TT	GA	AA	AG	TT	GG	GT	GG	GT	GG	GT	127481	
Qy	1981	GA	GA	GG	GA	CG	CT	GT	TT	GT	CT	TA	TA	TA	TA	TA	TA	TA	TA	TA	TA	TA	TA	TA	TA	2040	

Db	127480	GAG	AG	G	GA	CG	CG	CT	GT	TT	GT	GT	CT	TA	TA	TA	TA	TA	TA	TA	TA	TA	TA	TA	TA	127421																																																																																																																																																																																																																																																																																																																																																																																																																																																	
Qy	2041	TC	AA	TA	TA	C	AG	CA	GC	TT	TT	TA	T	TA	CT	TA	A	TT	T	TA	CT	TA	A	TT	T	TA	2100																																																																																																																																																																																																																																																																																																																																																																																																																																																
Db	127420	TC	AA	TA	TA	C	AG	CA	GC	TT	TT	TA	T	TA	CT	TA	A	TT	T	TA	CT	TA	A	TT	T	TA	127361																																																																																																																																																																																																																																																																																																																																																																																																																																																
Qy	2101	TC	AG	CA	GA	TT	CA	TA	CA	AA	GG	GG	CT	TT	CT	TA	AG	CG	CT	TA	AG	CG	CT	TA	AG	CG	2160																																																																																																																																																																																																																																																																																																																																																																																																																																																
Db	127360	TC	AG	CA	GA	TT	CA	TA	CA	AA	GG	GG	CT	TT	CT	TA	AG	CG	CT	TA	AG	CG	CT	TA	AG	CG	127301																																																																																																																																																																																																																																																																																																																																																																																																																																																
Qy	2161	TT	AT	CA	TT	CG	T	CC	CA	AG	AG	AG	CT	TA	GA	AG	A	TT	T	GA	AG	A	TT	T	GA	AG	2220																																																																																																																																																																																																																																																																																																																																																																																																																																																
Db	127300	TT	AT	CA	TT	CG	T	CC	CA	AG	AG	AG	CT	TA	GA	AG	A	TT	T	GA	AG	A	TT	T	GA	AG	127241																																																																																																																																																																																																																																																																																																																																																																																																																																																
Qy	2221	CT	CA	GG	GG	CT	GA	CC	CT	TA	TT	TA	GA	AA	CC	AA	GA	GG	GT	GG	TT	GA	AA	CC	TA	CT	CA	CG	2280																																																																																																																																																																																																																																																																																																																																																																																																																																														
Db	127240	CT	CA	GG	GG	CT	GA	CC	CT	TA	TT	TA	GA	AA	CC	AA	GA	GG	GT	GG	TT	GA	AA	CC	TA	CT	CA	CG	127181																																																																																																																																																																																																																																																																																																																																																																																																																																														
Qy	2281	TT	GA	TA	CC	AG	TC	GC	CA	CT	TC	GC	GG	AA	AA	AA	GG	GG	CT	TC	CC	AG	CC	AC	CC	CG	GA	TA	GG	2340																																																																																																																																																																																																																																																																																																																																																																																																																																													
Db	127180	TT	GA	TA	CC	AG	TC	GC	CA	CT	TC	GC	GG	AA	AA	AA	GG	GG	CT	TC	CC	AG	CC	AC	CC	CG	GA	TA	GG	127121																																																																																																																																																																																																																																																																																																																																																																																																																																													
Qy	2341	GG	GT	AA	GA	GA	AA	GA	GA	GG	CT	GG	GT	TA	GG	GG	CA	CT	GT	GT	TT	TA	AA	CA	GG	CA	CT	TC	2400																																																																																																																																																																																																																																																																																																																																																																																																																																														
Db	127120	GG	GT	AA	GA	GA	AA	GA	GA	GG	CT	GG	GT	TA	GG	GG	CA	CT	GT	GT	TT	TA	AA	CA	GG	CA	CT	TC	127061																																																																																																																																																																																																																																																																																																																																																																																																																																														
Qy	2401	TC	CT	CT	CG	GG	CG	CT	TA	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT	TT

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 216441)  
DOE Joint Genome Institute.  
Sequencing of Human Chromosome 19  
Unpublished  
2 (bases 1 to 216441)  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (30-MAR-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
On Jul 15, 2000 this sequence version replaced gi:7341654.  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: <http://www.jgi.doe.gov>  
-----  
Project Information  
Center Project Name: 726304\_B0691328  
Center clone name: CITB-H1\_2257C19  
-----  
Summary Statistics  
Consensus quality: 164680 bases at least Q40  
Consensus quality: 187929 bases at least Q30  
Consensus quality: 195340 bases at least Q20  
Estimated insert size: 159060; agarose-fp estimation  
Estimated insert size: 211311; sum-of-contigs estimation  
Quality coverage: 8.59 in Q20 bases; agarose-fp estimation  
Quality coverage: 6.46 in Q20 bases; sum-of-contigs estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 52 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 1041: contig of 1041 bp in length  
\* 1042 1141: gap of unknown length  
\* 1142 2370: contig of 1229 bp in length  
\* 2371 2470: gap of unknown length  
\* 2471 3552: contig of 1082 bp in length  
\* 3553 3652: gap of unknown length  
\* 3653 4683: contig of 1031 bp in length  
\* 4684 4783: gap of unknown length  
\* 4784 6178: contig of 1395 bp in length  
\* 6179 7442: gap of unknown length  
\* 7443 7542: gap of unknown length  
\* 7543 8661: contig of 1119 bp in length  
\* 8662 8761: gap of unknown length  
\* 8762 10030: contig of 1269 bp in length  
\* 10031 10130: gap of unknown length  
\* 10131 11491: contig of 1361 bp in length  
\* 11492 11591: gap of unknown length  
\* 11592 12967: contig of 1376 bp in length  
\* 12968 13067: gap of unknown length  
\* 13068 14173: contig of 1106 bp in length  
\* 14174 14273: gap of unknown length  
\* 14274 15840: contig of 1567 bp in length  
\* 15841 15940: gap of unknown length  
\* 15941 17242: contig of 1302 bp in length  
\* 17243 17342: gap of unknown length  
\* 17343 18449: contig of 1107 bp in length  
\* 18450 18549: gap of unknown length  
\* 18550 19683: contig of 1134 bp in length  
\* 19684 19783: gap of unknown length  
\* 19784 21247: contig of 1464 bp in length  
\* 21248 21347: gap of unknown length  
\* 21348 22952: contig of 1605 bp in length  
\* 22953 24552: gap of unknown length  
\* 24552: contig of 1400 bp in length

24453 24552: gap of unknown length  
24553 25825: contig of 1273 bp in length  
25826 25925: gap of unknown length  
25926 27109: contig of 1184 bp in length  
27110 27209: gap of unknown length  
27211 28292: contig of 1083 bp in length  
28293 28392: gap of unknown length  
28393 29744: contig of 1352 bp in length  
29745 29844: gap of unknown length  
29845 31042: contig of 1198 bp in length  
31043 31142: gap of unknown length  
31143 32629: contig of 1487 bp in length  
32630 32729: gap of unknown length  
32730 33698: contig of 1169 bp in length  
33699 33988: gap of unknown length  
33999 35704: contig of 1706 bp in length  
35705 35804: gap of unknown length  
35805 37471: contig of 1667 bp in length  
37472 37571: gap of unknown length  
37572 39285: contig of 1714 bp in length  
39286 39385: gap of unknown length  
39386 40577: contig of 1192 bp in length  
40578 40677: gap of unknown length  
40678 41930: contig of 1253 bp in length  
41931 42030: gap of unknown length  
42031 43743: contig of 1713 bp in length  
43744 43843: gap of unknown length  
43844 46054: contig of 2211 bp in length  
46055 46154: gap of unknown length  
46155 49379: contig of 3225 bp in length  
49380 49479: gap of unknown length  
49480 53371: contig of 3892 bp in length  
53372 53471: gap of unknown length  
53472 55756: contig of 2285 bp in length  
55757 55856: gap of unknown length  
55857 58019: contig of 2163 bp in length  
58020 58119: gap of unknown length  
58120 60327: contig of 2208 bp in length  
60328 60427: gap of unknown length  
60428 64223: contig of 3796 bp in length  
64224 64323: gap of unknown length  
64324 67218: contig of 2895 bp in length  
67219 67318: gap of unknown length  
67319 71797: contig of 4479 bp in length  
71798 71897: gap of unknown length  
71898 76856: contig of 4959 bp in length  
76857 76956: gap of unknown length  
76957 82060: contig of 5104 bp in length  
82061 82160: gap of unknown length  
82161 87557: contig of 5397 bp in length  
87558 87657: gap of unknown length  
87658 94226: contig of 6569 bp in length  
94227 94326: gap of unknown length  
94327 101878: contig of 7552 bp in length  
101879 101978: gap of unknown length  
101979 110629: contig of 8651 bp in length  
110630 110729: gap of unknown length  
110730 119083: contig of 8354 bp in length  
119084 119183: gap of unknown length  
119184 128919: contig of 9736 bp in length  
128920 129019: gap of unknown length  
129020 142791: contig of 13772 bp in length  
142792 142891: gap of unknown length  
142892 156079: contig of 13188 bp in length  
156080 156179: gap of unknown length  
156180 183092: contig of 26913 bp in length  
183093 216441: contig of 33249 bp in length.  
Location/Qualifiers  
1.216441  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"



```
/chromosome="19"
/clone="CTD-2257C19"
/clone_lib="CalTech human BAC library D"
gap 1042.1141
gap /estimated_length=unknown 2371.2470
gap /estimated_length=unknown 3553.3652
gap /estimated_length=unknown 4684.4783
gap /estimated_length=unknown 6179.6278
gap /estimated_length=unknown 7443.7542
gap /estimated_length=unknown 8662.8761
gap /estimated_length=unknown 10031.10130
gap /estimated_length=unknown 11492.11591
gap /estimated_length=unknown 12968.13067
gap /estimated_length=unknown 14174.14273
gap /estimated_length=unknown 15841.15940
gap /estimated_length=unknown 17243.17342
gap /estimated_length=unknown 18450.18549
gap /estimated_length=unknown 19684.19783
gap /estimated_length=unknown 21248.21347
gap /estimated_length=unknown 22953.23052
gap /estimated_length=unknown 24453.24552
gap /estimated_length=unknown 25826.25925
gap /estimated_length=unknown 27110.27209
gap /estimated_length=unknown 28293.28392
gap /estimated_length=unknown 29745.29844

Query Match 92.7%; Score 2892.8; DB 14; Length 216441;
Best Local Similarity 99.9%; Pred. No. 0;
Matches 2894; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ACTAGAGGTGGGGTTTGCCGTTGGAGCACCACCAAGTGAAGCGCAACCGCGCAGGAGC 60
DB 117034 ACTAGAGGTGGGGTTTGCCGTTGGAGCACCACCAAGTGAAGCGCAACCGCGCAGGAGC 116975
QY 61 ACCTGACCCCGGCGGCGCCAGCCCTCGGATTGGCAAGTCACTGCTGCTTTGGGGCAGC 120
DB 116974 ACCTGACCCCGGCGGCGCCAGCCCTCGGATTGGCAAGTCACTGCTGCTTTGGGGCAGC 116915
QY 121 GAGGTGCCAGTCTTCCGCGGACCCGAGCTCTGTCCGACAGAGGTCCGGAGTCACT 180
DB 116914 GAGGTGCCAGTCTTCCGCGGACCCGAGCTCTGTCCGACAGAGGTCCGGAGTCACT 116855
QY 181 ATAGCTGGGTTCTAATGCCATCAAGCAAGCAAACTCCGCGGAGCTCGCTTTTAA 240
DB 116854 ATAGCTGGGTTCTAATGCCATCAAGCAAGCAAACTCCGCGGAGCTCGCTTTTAA 116795
QY 241 CCTGGGCTCAGTTTCCCATCGTAAATAGAACGGGTTGATCTCCGAGGCGTAAAC 300
DB 116794 CCTGGGCTCAGTTTCCCATCGTAAATAGAACGGGTTGATCTCCGAGGCGTAAAC 116735
QY 301 TTCAGAACTCGATGGGCGAAGGAGGAGGATGGGCCAACCAAGTGAAGCTCC 360
```

```
DB 116734 TTCAGAACTCGATGGGCGAAGGAGGAGATGGGCCAACCAAGTGAAGCTCC 116675
QY 361 CGCGTGAAGCCCGGCTTACCACTGATCCAGGGGGTGGCAAGTCCCGGCGGAGACGAGCGGG 420
DB 116674 CGCGTGAAGCCCGGCTTACCACTGATCCAGGGGGTGGCAAGTCCCGGCGGAGACGAGCGGG 116615
QY 421 GTGGCGGGGCTCTAAGAAACCTTACCCGCGCGCTTTGGCAGCGCTTAAAGCGAGCGCG 480
DB 116614 GTGGCGGGGCTCTAAGAAACCTTACCCGCGCGCTTTGGCAGCGCTTAAAGCGAGCGCG 116555
QY 481 CGGCTCTAGCGCTTGTCCCGGAGTTTGACCCACGAGAGATGGGAGCCGACCTTC 540
DB 116554 CGGCTCTAGCGCTTGTCCCGGAGTTTGACCCACGAGAGATGGGAGCCGACCTTC 116495
QY 541 AGCTTCGAGGAGAGCCACCTGTGAGAGCCAGGGCGGTGACAGACACGACGTGTGACTGG 600
DB 116494 AGCTTCGAGGAGAGCCACCTGTGAGAGCCAGGGCGGTGACAGACACGACGTGTGACTGG 116435
QY 601 AGTGGCGCTGGGGAGATGAGACGAGGAGCGGGAGACGCTTACCGAGGCTCCCTCTGCGC 660
DB 116434 AGTGGCGCTGGGGAGATGAGACGAGGAGCGGGAGACGCTTACCGAGGCTCCCTCTGCGC 116375
QY 661 GCCCGCTCCGACGAGGCGCACGTCGAGGAGTCCCGGACGAGCTCCGTGACGTTGCGGTA 720
DB 116374 GCCCGCTCCGACGAGGCGCACGTCGAGGAGTCCCGGACGAGCTCCGTGACGTTGCGGTA 116315
QY 721 GCGCCGACGAGTCAACGACACATGAAAGCGTTCGTCGCGCGGCCCAAGGCGGAGATG 780
DB 116314 GCGCCGACGAGTCAACGACACATGAAAGCGTTCGTCGCGCGGCCCAAGGCGGAGATG 116255
QY 781 GGGGTTAAGCAATCTCCGCGCTGAGGGGAGGCTTAAAGGCGCGGCGCGCGCGCGCC 840
DB 116254 GGGGTTAAGCAATCTCCGCGCTGAGGGGAGGCTTAAAGGCGCGGCGCGCGCGCGCC 116195
QY 841 AGCCGAGCCCAACCGCATGAGGAGGAGAGAGTGCAGAGCGCTGCTGACAGGCTCAACA 900
DB 116194 AGCCGAGCCCAACCGCATGAGGAGGAGAGAGTGCAGAGCGCTGCTGACAGGCTCAACA 116135
QY 901 AGACGACTGCGGTGATCAACCAACCTGATGCTGACCGTGGTGGCTCGGCGGACCTGCGAGA 960
DB 116134 AGACGACTGCGGTGATCAACCAACCTGATGCTGACCGTGGTGGCTCGGCGGACCTGCGAGA 116075
QY 961 ACCTCGGACGAGGCTGCAAAAGACGCGCACAGAGCGCACAGAGCTGACGATGCTCACT 1020
DB 116074 ACCTCGGACGAGGCTGCAAAAGACGCGCACAGAGCGCACAGAGCTGACGATGCTCACT 116015
QY 1021 GCGCCCGCTGACTGCTGTGCTGCGCGACCGGGGCTTGCGCGACGAGCGCGCGAGT 1080
DB 116014 GCGCCCGCTGACTGCTGTGCTGCGCGACCGGGGCTTGCGCGACGAGCGCGCGAGT 115955
QY 1081 TCGAGCGGCTCGGGTGGGCTTCTCGAGGCTGCTGAGACCTGTGAGAGCGGACATGCGAC 1140
DB 115954 TCGAGCGGCTCGGGTGGGCTTCTCGAGGCTGCTGAGACCTGTGAGAGCGGACATGCGAC 115895
QY 1141 GCTCGCTGAGCTGAGCGCGCGCTTCCGCTTGCACGCGCGCGCGAGACCGCTGATGCGCA 1200
DB 115894 GCTCGCTGAGCTGAGCGCGCGCTTCCGCTTGCACGCGCGCGCGAGACCGCTGATGCGCA 115835
QY 1201 CAGGTGTGGCTGGCGCTCTCTCCGCGTGGCGCGCGCGCTTGAGCACCGCGACCTTGC 1260
DB 115834 CAGGTGTGGCTGGCGCTCTCTCCGCGTGGCGCGCGCGCTTGAGCACCGCGACCTTGC 115775
QY 1261 GGTGAGGCGGAGGCGGACCTTGCAGTTCGAGGACCTGCGGAGGCTGAGGCGGAGTCC 1320
DB 115774 GGTGAGGCGGAGGCGGACCTTGCAGTTCGAGGACCTGCGGAGGCTGAGGCGGAGTCC 115715
QY 1321 TTCAAGTGGCGAGATGATGCAACAATGAGATGAAGTCAACGATGCCCGCTGAGACCG 1380
DB 115714 TTCAAGTGGCGAGATGATGCAACAATGAGATGAAGTCAACGATGCCCGCTGAGACCG 115655
QY 1381 TGCAGACCCGCGAGCGCGCGCGCGAGCTCTGTCCAAGGTCAAGCGCGCGCTCTCT 1440
DB 115654 TGCAGACCCGCGAGCGCGCGCGCGCGAGCTCTGTCCAAGGTCAAGCGCGCGCTCTCT 115595
```



QY	1441	CGGTGTGTCTTTGACAGAGCCGGGGGGGTTCGACCACGAGAAAGCCCTGGCCGCA	1500
Db	11594	CGGTGTGTCTTTGACAGAGCCGGGGGGGTTCGACCACGAGAAAGCCCTGGCCGCA	115935
QY	1501	TCCTTTGGGGCCCGTGTGTGGCGGTGGCCCTTACCCGTGTGGGGAGACTGA	1560
Db	115534	TCCTTTGGGGCCCGTGTGTGGCGGTGGCCCTTACCCGTGTGGGGAGACTGA	115475
QY	1561	GCTGACAGACACCCGACGCGCGCTGTGCGCTCCCTCCCTGAGAAAAGACTCG	1620
Db	115474	GCTGACAGACACCCGACGCGCGCTGTGCGCTCCCTCCCTGAGAAAAGACTCG	115415
QY	1621	GATGGGTGTGGGCTGTGGCCTGTGCAAGGGAGTGTCTTAAACCCCGTGTGCATGG	1680
Db	115414	GATGGGTGTGGGCTGTGGCCTGTGCAAGGGAGTGTCTTAAACCCCGTGTGCATGG	115355
QY	1681	GTACACGCGCGTTTCCAGTGCAATGTGCTGGGACAGACAGTTTTCTCTGTGTGGC	1740
Db	115354	GTACACGCGCGTTTCCAGTGCAATGTGCTGGGACAGACAGTTTTCTCTGTGTGGC	115295
QY	1741	CCGGAGAAAGTTAACTTTGCGCGCGCGCTCAGGGCAATACCGTTAAGTCTGAGAGCT	1800
Db	115294	CCGGAGAAAGTTAACTTTGCGCGCGCGCTCAGGGCAATACCGTTAAGTCTGAGAGCT	115235
QY	1801	TTATTCCCTAATTAAGAAAACGTCACAGTGACCTTAGATCCCTCGAGTTAATGAGTT	1860
Db	115234	TTATTCCCTAATTAAGAAAACGTCACAGTGACCTTAGATCCCTCGAGTTAATGAGTT	115175
QY	1861	AAACAATGTGTGTGGGGCTTTTACAGGAGTCCGAGTTGGTGGCCACCCCTGCCA	1920
Db	115174	AAACAATGTGTGTGGGGCTTTTACAGGAGTCCGAGTTGGTGGCCACCCCTGCCA	115115
QY	1921	GGGTGGCCCCCTTTCGCGTGGGACAGTTGAAAAGTGGGGTGGAGTAAGTTTG	1980
Db	115114	GGGTGGCCCCCTTTCGCGTGGGACAGTTGAAAAGTGGGGTGGAGTAAGTTTG	115055
QY	1981	GAGAGGACGCTGTTGTTCTATGTGTGTCTGTCTCCGACAAAGAAAATTGCA	2040
Db	115054	GAGAGGACGCTGTTGTTCTATGTGTGTCTGTCTCCGACAAAGAAAATTGCA	114995
QY	2041	TCAAAATGCAGACGTTTTATTACTTAATCTTTCAGGGCTTAAATTTAGAGATGTCC	2100
Db	114994	TCAAAATGCAGACGTTTTATTACTTAATCTTTCAGGGCTTAAATTTAGAGATGTCC	114935
QY	2101	TGAGAGAGCTTCAACAAGGGCTTCTCTAAGACGCGCTACAGCCCTCTAGCAGAGT	2160
Db	114934	TGAGAGAGCTTCAACAAGGGCTTCTCTAAGACGCGCTACAGCCCTCTAGCAGAGT	114875
QY	2161	TTATCATTCGTCCCAAGAGCAGCTAGAAAGATTGAGTCAATACCTCCACTGGCG	2220
Db	114874	TTATCATTCGTCCCAAGAGCAGCTAGAAAGATTGAGTCAATACCTCCACTGGCG	114815
QY	2221	CTCAGGGGCTGACCTATTATTAGAAAACAAGAGGGTGGGTGAACCTACTCACGGAC	2280
Db	114814	CTCAGGGGCTGACCTATTATTAGAAAACAAGAGGGTGGGTGAACCTACTCACGGAC	114755
QY	2281	TTGGAATCCAGTGCACACTTGTCTCGGAAAAAGGGCTCTCCCAAGCACCGGAATGG	2340
Db	114754	TTGGAATCCAGTGCACACTTGTCTCGGAAAAAGGGCTCTCCCAAGCACCGGAATGG	114695
QY	2341	GGGTAGAGAAAGACAGAGGCTTGGGGTAAAGGCACTGTGTTTAAACAGCACTTTC	2400
Db	114694	GGGTAGAGAAAGACAGAGGCTTGGGGTAAAGGCACTGTGTGTAAACAGCACTTTC	114635
QY	2401	TCCTTCTCGGGGCTTAATTTTGTTCAGAACTAGACCAAGTGTGAACTCTCTTGGCA	2460
Db	114634	TCCTTCTCGGGGCTTAATTTTGTTCAGAACTAGACCAAGTGTGAACTCTCTTGGCA	114575
QY	2461	GGAGGGCTGGGAATCTCTTTTAGAGACTTAATCTTAATTTATCCCTCGGAATGTGCTGC	2520
Db	114574	GGAGGGCTGGGAATCTCTTTTAGAGACTTAATCTTAATTTATCCCTCGGAATGTGCTGC	114515

QY	2521	TGGCCAGTAGAGAGGGCTGGCTTTGGCACTGCCCTACCCCGCGCTGCCGCTCCCGG	2580
Db	114514	TGGCCAGTAGAGAGGGCTGGCTTTGGCACTGCCCTACCCCGCGCTGCCGCTCCCGG	114455
QY	2581	GGTAATGTGGCACTTACTGGCCACAGAGGTTTTGAGCCAAATCAGCTCTGAGCTGGGTTA	2640
Db	114454	GGTAATGTGGCACTTACTGGCCACAGAGGTTTTGAGCCAAATCAGCTCTGAGACTGGGTTA	114395
QY	2641	GAATGTAA CAGCTTTAACTTGGGATTTAAGAAAGCTTTTAAAGGTAAATATCCTTGAAA	2700
Db	114394	GAATGTAA CAGCTTTAACTTGGGATTTAAGAAAGCTTTTAAAGGTAAATATCCTTGAAA	114335
QY	2701	GAATAATACGTAAACACAGCGCTGTACATAGAAACCTGTATTTTAAATAAGAACGCTGG	2760
Db	114334	GAATAATACGTAAACACAGCGCTGTACATAGAAACCTGTATTTTAAATAAGAACGCTGG	114275
QY	2761	GCCATGAACCTCATACCTGCGCAATGAGTCMAACATAGTATCTTATGTATGACTTATGATT	2820
Db	114274	GCCATGAACCTCATACCTGCGCAATGAGTCMAACATAGTATCTTATGTATGACTTATGATT	114215
QY	2821	ACTAATAATATATTTTCATCTACTTGTGAAGTTGATAGTCTTCCCGCCCGCCCACTTTT	2880
Db	114214	ACTAATAATATATTTTCATCTACTTGTGAAGTTGATAGTCTTCCCGCCCGCCCACTTTT	114155
QY	2881	TCTTTTTTGAAGCAGG 2896	
Db	114154	TCTTTTTTGAAGCAGG 114139	

RESULT 5	18615 bp	DNA	linear	PRI 13-JUL-2002
AC008474				
LOCUS	AC008474			
DEFINITION	Homo sapiens chromosome 19 clone CTC-379B2, complete sequence.			

VERSION AC008474.9 GI:21743729  
KEYWORDS HTG.

SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 186115)

**AUTHORS** DOE Joint Genome Institute and Stanford Human Genome Center  
**TITLE** Direct Submission

**JOURNAL** Unpublished  
**REFERENCE** 2 (bases 1 to 186115)

<b>AUTHORS</b>	<b>DOE Joint Genome Institute.</b>
<b>TITLE</b>	<b>Direct Submission</b>

**JOURNAL** Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA

REFERENCE 3 (bases 1 to 186115)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center

**TITLE** Direct Submission  
**JOURNAL** Submitted (01-JUL-2000) DOE Joint Genome Institute, 2800 Mitchell

4 (bases 1 to 186115)  
Drive, Walnut Creek, CA 94598, USA

**AUTHORS** DOE Joint Genome Institute and Stanford Human Genome Center.  
**TITLE** Direct Submission

JOURNAL  
Submitted (17-OCT-2001) DOB Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA

REFERENCE 5 (bases 1 to 186115)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.

**TITLE** Direct Submission  
**JOURNAL** Submitted (13-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell

COMMENT On Jul 13, 2002 this sequence replaced gi:16197759.  
Drive, Walnut Creek, CA 94598, USA

Pratt Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov

Finishing completed at Stanford Human Genome Center  
www-shgc.stanford.edu

quality: phrap quality >=40 99.8% of sequence,  
Estimated Total Number of Errors is 0.4.

## FEATURES

### Location/Qualifiers

```
source 1.186115
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/clone="CTC-379B2"

ORIGIN

Query Match 91.6%; Score 2860.8; DB 8; Length 186115;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2887; Conservative 0; Mismatches 7; Indels 2; Gaps 2;

QY 1 ACTAGAGGTGGGGTTGACCTTGGAAGCTCCGACCAAGTGAAGCGCAAGCGGCAAGGAGC 60
DB 180657 ACTAGAGGTGGGGTTGACCTTGGAAGCAACCGACCAAGTGAAGCGCAAGCGGCAAGGAGC 180716
QY 61 ACCTGACCCCGGCGGCGCCAGCGCCCTCGGATTTGACAGTCACTGCTCGCTTTGGGCAAG 120
DB 180717 ACTTGACCCCGGCGGCGCCAGCGCCCTCGGATTTGACAGTCACTGCTCGCTTTGGGCAAG 180776
QY 121 GAGGTGCCAGATCTTCGCGGAGCAACCGACGATCTGTGCGCGACAGGGTCCGGAGTCAAT 180
DB 180777 GAGGTGCCAGATCTTCGCGGAGCAACCGACGATCTGTGCGCGACAGGGTCCGGAGTCAAT 180836
QY 181 ATAGCTGGGTTGATGTCCTGACAGGCAAGGCAAAAATCTCGGGGAGCGCTGGCCGCTTTTA 240
DB 180837 ATAGCTGGGTTGATGTCCTGACAGGCAAGGCAAAAATCTCGGGGAGCGCTGGCCGCTTTTA 180896
QY 241 CCTGGGCTCAAGTTTCCCATCCGTAATAAGACGGATTGATCTCCGAGCGCTTACA 300
DB 180897 CCTGGGCTCAAGTTTCCCATCCGTAATAAGACGGATTGATCTCCGAGGCTTACA 180956
QY 301 TTCCAGAACTTCGATATGGGCGAAGGGAGAGGATATGGGCGCACCCACATGATCTTCCC 360
DB 180957 TTCCAGAACTTCGATATGGGCGAAGGGAGAGGATATGGGCGCACCCACATGATCTTCCC 181016
QY 361 CGGATGAGACCCCGCTTACCACTGATCAAGGGGTATGGAGCTCCGAGCGGAGCAAGCGAG 420
DB 181017 CGGATGAGACCCCGCTTACCACTGATCAAGGGGTATGGAGCTCCGAGCGGAGCAAGCGAG 181076
QY 421 GTGGGCGGGTCTTAGAGAAACCTTACCAGCGCCCTTTGGCAGCGCCTTAAGCGGAGCGCG 480
DB 181077 GTGGGCGGGTCTTAGAGAAACCTTACCAGCGCCCTTTGGCAGCGCCTTAAGCGGAGCGCG 181136
QY 481 CGGCTCTGACGCTGCTTCCCGCGAGTTGGCACCCACGAGATATGGGACCGCACCTTC 540
DB 181137 CGGCTCTGACGCTGCTTCCCGCGAGTTGGCACCCACGAGATATGGGACCGCACCTTC 181196
QY 541 AGCTTGACAGGAGCAACCGTGAAGGCAAGGCGGATGCAAGACAGAGCTGATCCG 600
DB 181197 AGCTTGACAGGAGCAACCGTGAAGGCAAGGCGGATGCAAGACAGAGCTGATCCG 181256
QY 601 AGTGCGCTTGAGGAGATGAGACGAGGAGAGCGGGAGCCGCTTAAGGGGCTCCCTTGGCG 660
DB 181257 AGTGCGCTTGAGGAGATGAGACGAGGAGAGCGGGAGCCGCTTAAGGGGCTCCCTTGGCG 181316
QY 661 GCCCCTCCGCGAGAGCGACATCTGAGGGTCCCGGGCGGGCTCCGTGACATTTGGGGGTA 720
DB 181317 GCCCCTCCGCGAGAGCGACATCTGAGGGTCCCGGGCGGGCTCCGTGACATTTGGGGGTA 181376
QY 721 GCGCCGAGGAGTCAAGGACATGAAGAGGATTCGTCGCGGCGGCGCAAGGCGGGGATG 780
DB 181377 GCGCCGAGGAGTCAAGGACATGAAGAGGATTCGTCGCGGCGGCGCAAGGCGGGGATG 181436
QY 781 GGGGTTAGCCATCTCTGCGCGCTGAGGGGAGGGTTAACGGGCGCGGCGCGGCGCC 840
DB 181437 GGGGTTAGCCATCTCTGCGCGCTGAGGGGAGGGTTAACGGGCGCGGCGCGGCGCC 181496
QY 841 AGCCGAGAGCCACCGGAGATGAGGAGAGGAGATGACAGGCGCTGCTGACCGGCTTAA 900
DB 181497 AGCCGAGAGCCACCGGAGATGAGGAGAGGAGATGACAGGCGCTGCTGACCGGCTTAA 181556
QY 901 AGACGACTGCTGCTACACCACTGTGCTGACCGTGGTGGCTCGGCGGACTCGAGA 960
```

```
DB 181557 AAGACATGCTGTCTACTACCACTGTGTCTACCGTGGTGGCTCTGGGAGACTGGCAGA 181616
QY 961 ACCTCGAGGAGAGTGTCAAAAAGACCGCGCAGAAAGCGCAGAGAGTGTGCACCT 1020
DB 181617 ACCTCGAGGAGAGTGTCAAAAAGACCGCGCAGAAAGCGCAGAGAGTGTGCACCT 181676
QY 1021 GGGCCCGGTGACTGTGTGTGCGCGACCGGGGCTGGCCCGGACGAGCGGCGAGT 1080
DB 181677 GGGCCCGGTGACTGTGTGTGCGCGACCGGGGCTGGCCCGGACGAGCGGCGAGT 181736
QY 1081 TCGAGCGGCTTGGGTGGGCTTCTGCGGCTGTGGAACCTGTGGAAGCGGACATGCGAC 1140
DB 181737 TCGAGCGGCTTGGGTGGGCTTCTGCGGCTGTGGAACCTGTGGAAGCGGACATGCGAC 181796
QY 1141 GCTGCTGAGAGTGGAGCGCGCGTTCCGCTGACGCGCGCGGCGACCGCTGTGCGCA 1200
DB 181797 GCTGCTGAGAGTGGAGCGCGCGTTCCGCTGACGCGCGCGGCGACCGCTGTGCGCA 181856
QY 1201 CAGGTGTGCTGGCGCTTCTCCGCGGTGAGCGCGCGGCTGTGAGCAACCGGACCTGC 1260
DB 181857 CAGGTGTGCTGGCGCTTCTCCGCGGTGAGCGCGCGGCTGTGAGCAACCGGACCTGC 181916
QY 1261 GCGTGAAGCGGAGGAGCGACTTGACAGTTCGCGGACCTGCGGAGCTGAGCGGAGTTC 1320
DB 181917 GCGTGAAGCGGAGGAGCGACTTGACAGTTCGCGGACCTGCGGAGCTGAGCGGAGTTC 181976
QY 1321 TTCAAGTGGGCGAGATGATGCAACATGAGATGAAGGTCAACGTGCCCGCTGACCG 1380
DB 181977 TTCAAGTGGGCGAGATGATGCAACATGAGATGAAGGTCAACGTGCCCGCTGACCG 182036
QY 1381 TGCAGGCCGCGAGCGCGCGCGCGCGCTCTGCAACCGTCAAGCGCGCGCCCTCT 1440
DB 182037 TGCAGGCCGCGAGCGCGCGCGCGCGCTCTGCAACCGTCAAGCGCGCGCCCTCT 182096
QY 1441 CGGTGCTGCTTGGAGAGCGCGGGGGGATTGCGACCCCGAGAAAGCGCTGCGCGCA 1500
DB 182097 CGGTGCTGCTTGGAGAGCGCGGGGGGATTGCGACCCCGAGAAAGCGCTGCGCGCA 182156
QY 1501 TCCCTTTTGGCGCGCTGTGCTGCGCGCTGTGAGCCCTTACCGTGTGTGCGAAAGCTGA 1560
DB 182157 TCCCTTTTGGCGCGCTGTGCTGCGCGCTGTGAGCCCTTACCGTGTGTGCGAAAGCTGA 182216
QY 1561 GCTGACAGACACCGACCGCGCGCTGTGCGCGCTGTGAGCCCTTACCGTGTGTGCGAAAGCTGA 1620
DB 182217 GCTGACAGACACCGACCGCGCGCTGTGCGCGCTGTGAGCCCTTACCGTGTGTGCGAAAGCTGA 182276
QY 1621 GATGGGTGTGGGATCTGGCTGTGCAAGGAGAGTGTCTTAACCCCGTGTGTGCAATG 1680
DB 182277 GATGGGTGTGGGATCTGGCTGTGCAAGGAGAGTGTCTTAACCCCGTGTGTGCAATG 182336
QY 1681 GTACACGCGCGCTTTCAGATGCAATCTGCTGAGGACAGACAGGTTTCTCTTGTGCGC 1740
DB 182337 GTACACGCGCGCTTTCAGATGCAATCTGCTGAGGACAGACAGGTTTCTCTTGTGCGC 182396
QY 1741 CGGGAGAGATTAACCTTTGGCGCGCGCGCTGAGGGGATTAACGCTTAACGCTGACAGAGCT 1800
DB 182396 CGGGAGAGATTAACCTTTGGCGCGCGCGCTGAGGGGATTAACGCTTAACGCTGACAGAGCT 182455
QY 1801 TTATTCCTATTATTAAGAAACCGTCAAGTGAACCTAGATCCCTCCGATTATGAGTT 1860
DB 182456 TTATTCCTATTATTAAGAAACCGTCAAGTGAACCTAGATCCCTCCGATTATGAGTT 182515
QY 1861 AACACATGTGCTGTGGGGCGCTTTTACAGGAGATCCGAGTTCCGATCCACCTTGC 1920
DB 182516 AACACATGTGCTGTGGGGCGCTTTTACAGGAGATCCGAGTTCCGATCCACCTTGC 182575
QY 1921 GGGTGGCCCGCTTTCGCGTGGGACAGTTTGAAGATGAGGTGGGATGGAAGTTTG 1980
DB 182576 GGGTGGCCCGCTTTCGCGTGGGACAGTTTGAAGATGAGGTGGGATGGAAGTTTG 182635
QY 1981 GAGAGGAGCGCTGTTGTTGTTATGTGTGTGTTCCGCGACAAAGAAATTTGCA 2040
```

Db 182636 GAGAGGACGCTTTGGTTCATATGATGTGTCTCTGTTCCGGACAAAGAAAATTGCAG 182695  
 QY 2041 TCAAAATGTCAGGACCTTTTAACTTAACTTTTCAAGGCTTAAATTTAGAGAGTGC 2100  
 Db 182696 TCAAAATGTCAGGACCTTTTAACTTAACTTTTCAAGGCTTAAATTTAGAGAGTGC 182755  
 QY 2101 TGAGACAGTTCATACAAAGGCTTCTCTAAGACGCTACAGCCCTTCTAGACAGT 2160  
 Db 182756 CGAGACAGTTCATACAAAGGCTTCTCTAAGACGCTACAGCCCTTCTAGACAGT 182815  
 QY 2161 TTAATCATTTGTCCTCCCAAGACAGCTAGAAAGATTTAGAGTATGACCTCCCATGCGC 2220  
 Db 182816 TTAATCATTTGTCCTCCCAAGACAGCTAGAAAGATTTAGAGTATGACCTCCCATGCGC 182875  
 QY 2221 CTCAGGGGCTGACCTTAACTTTAGAAACCAAGAGGCTTGAACCTTACTCTACAGAC 2280  
 Db 182876 CTCAGGGGCTGACCTTAACTTTAGAAACCAAGAGGCTTGAACCTTACTCTACAGAC 182935  
 QY 2281 TTGATTCAGTGGGCACTTTCCTGCGAAAAGGCTCTCCCAAGCAACCCGAGATGG 2340  
 Db 182936 TTGATTCAGTGGGCACTTTCCTGCGAAAAGGCTCTCCCAAGCAACCCGAGATGG 182995  
 QY 2341 GGGTAAAGAGAAAGACAGAGGCTTGGGTAAGGCACTGTGTTTAAACAGGCACTTTC 2400  
 Db 182996 GGGTAAAGAGAAAGACAGAGGCTTGGGTAAGGCACTGTGTTTAAACAGGCACTTTC 183055  
 QY 2401 TCTTTCTCTGGGCTTATTTTGTTCAGAACTAGACCAAGGTGTGAACCTCTTTCGA 2460  
 Db 183056 TCTTTCTCTGGGCTTATTTTGTTCAGAACTAGACCAAGGTGTGAACCTCTTTCGA 183115  
 QY 2461 GGAGGCTGGGAATCCCTTTAGAGCACTTAATCTTATTAATCCCGGAATGGGCTG 2520  
 Db 183116 GGAGGCTGGGAATCCCTTTAGAGCACTTAATCTTATTAATCCCGGAATGGGCTG 183175  
 QY 2521 TGGCAGTAGAGAGGCTGCTTGGGAGACTCCCTGAACCCCGGCTGCCCCCTCCG 2580  
 Db 183176 TGGCAGTAGAGAGGCTGCTTGGGAGACTCCCTGAACCCCGGCTGCCCCCTCCG 183235  
 QY 2581 GGTAAATGTGCACTTACTGCGCCCAAGAGTTTGAAGCAATGACTCTGAGACTGGGTTA 2640  
 Db 183236 GGTAAATGTGCACTTACTGCGCCCAAGAGTTTGAAGCAATGACTCTGAGACTGGGTTA 183295  
 QY 2641 GAATGTAACGCTTAACTTGGGATTTAAGAACTTTTAAAGSTATAATCTCTGAAA 2700  
 Db 183296 GAATGTAACGCTTAACTTGGGATTTAAGAACTTTTAAAGSTATAATCTCTGAAA 183355  
 QY 2701 GAAAAATGACGTAAACCAAGCGGTATCTATGAAGCTGTATTTTAAAGAAACGCTGG 2760  
 Db 183356 GAAAAATGACGTAAACCAAGCGGTATCTATGAAGCTGTATTTTAAAGAAACGCTGG 183415  
 QY 2761 GCCATGAATCATACCTGCCAATGAGTCAAACTAGTATTTTATGTAATCTTGAAT 2820  
 Db 183416 GCCATGAATCATACCTGCCAATGAGTCAAACTAGTATTTTATGTAATCTTGAAT 183475  
 QY 2821 ACTAAATATATATTTTCACTACTCTGAAGTATGATCTTCCCCCCCCCACTTTT 2880  
 Db 183476 ACTAAATATATATTTTCACTACTCTGAAGTATGATCTTCCCCCCCCCACTTTT 183534  
 QY 2881 TCTTTTGGAGGAG 2896  
 Db 183535 TCTTTTGGAGGAG 183550

RESULT 6  
 CQ850499 2327 bp DNA linear PAT 23-AUG-2004  
 LOCUS CQ850499 Sequence 968 from Patent EP1447413.  
 DEFINITION CQ850499  
 ACCESSION CQ850499  
 VERSION CQ850499.1 GI:51508711  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 REFERENCE  
 AUTHORS Isegai, T., Yamamoto, J., Nishikawa, T., Isono, Y., Sugiyama, T.,  
 Otsuki, T., Wakematsu, A., Ishii, S., Nagai, K. and Irie, R.  
 TITLE Full-length human CDNA  
 JOURNAL Patent: EP 1447413-A 968 18-AUG-2004;  
 Research Association for Biotechnology (JP)  
 FEATURES  
 source  
 1..2327  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 ORIGIN  
 Query Match 73.9%; Score 2308; DB 6; Length 2327;  
 Best Local Similarity 99.7%; Pred. No. 4,4e-290;  
 Matches 2322; Conservative 0; Mismatches 5; Indels 1; Gaps 1;  
 QY 444 ACCCGGCGCCCTTGGCAGCGCTTAAGGCGAGCGGCGCTGAGGCTGCTGCCCC 503  
 Db 1 ACCCGGCGCCCTTGGCAGCGCTTAAGGCGAGCGGCGCTGAGGCTGCTGCCCC 60  
 QY 504 GGAGTTGGACCCACGAGAGATGGGACCGCACCTTCACTTGCAGGAGCCACCGTGG 563  
 Db 61 GGAGTTGGACCCACGAGAGATGGGACCGCACCTTCACTTGCAGGAGCCACCGTGG 120  
 QY 564 AGGCCAGGCGGTGCAGAGACACAGCTGTGACTCGAGTGGCTGGGAGATGAGACG 623  
 Db 121 AGGCCAGGCGGTGCAGAGACACAGCTGTGACTCGAGTGGCTGGGAGATGAGACG 180  
 QY 624 AGGAGCGGGGAGACCGCTAACCGGGCTCCCTCTGCGCGCCCGTCCGACAGGCGGACGT 663  
 Db 181 AGGAGCGGGGAGACCGCTAACCGGGCTCCCTCTGCGCGCCCGTCCGACAGGCGGACGT 240  
 QY 684 CGAGGCTCCGCGCGGCTCCGTGACGTTGGCGGTAGAGCGGACGAGTCAACGACAT 743  
 Db 241 CGAGGCTCCGCGCGGCTCCGTGACGTTGGCGGTAGAGCGGACGAGTCAACGACAT 300  
 QY 744 GAAAGCGTTTCGTCGCGCGCGCCCAAGGCGGAGATGGGAGTTAGCAATCTTGC 803  
 Db 301 GAAAGCGTTTCGTCGCGCGCGCCCAAGGCGGAGATGGGAGTTAGCAATCTTGC 360  
 QY 804 CTGAGGGGAGGCTAACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 863  
 Db 361 CTGAGGGGAGGCTAACCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420  
 QY 864 AGGAGAGATGACAGCGCTGCTGACAGGCTCAACAAAGCACTGCTGCTACCAAC 923  
 Db 421 AGGAGAGATGACAGCGCGCTGCTGACAGGCTCAACAAAGCACTGCTGCTACCAAC 480  
 QY 924 CTGATGTCACCGTGTGGCTTGGCGGACCTGCGAGACCTTGGCGGAGAGCTGCAAAAG 983  
 Db 481 CTGATGTCACCGTGTGGCTTGGCGGACCTGCGAGACCTTGGCGGAGAGCTGCAAAAG 540  
 QY 984 AGCGCGCAAGAGCGGAGAGGCTGCGGCTGCTGCACTTGGCGCGCGCGCGCGCGCG 1043  
 Db 541 AGCGCGCAAGAGCGGAGAGGCTGCGGCTGCTGCACTTGGCGCGCGCGCGCGCGCG 600  
 QY 1044 CGCGACCGGGGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1103  
 Db 601 CGCGACCGGGGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 660  
 QY 1104 TCGGCTGCTGAGACCTGCTGGAAGGAGCAATGCGACCGCGCTGAGCTGGGCGCGCG 1163  
 Db 661 TCGGCTGCTGAGACCTGCTGGAAGGAGCAATGCGACCGCGCTGAGCTGGGCGCGCG 720  
 QY 1164 TTCCCGCTGCAAGCG 1223  
 Db 721 TTCCCGCTGCAAGCG 780  
 QY 1224 GCGTGGCGCGCGCGCGCTGAGCAACCGCAAGCTTGGCGCTGAGCGGAGCGGACCTTC 1283

Db 781 GGGGTGGGGGCGGCGGCTGAGCACCGGACGCTGGGCTGAGGCGGAGCGACCTTC 840  
Qy 1284 GAGGTGCGGAGACTGCGGAGCTGAGCGGAGGTCCTTCAAGTGGGCGAGATGATGAC 1343  
Db 841 GAGGTGCGGAGACTGCGGAGCTGAGCGGAGGTCCTTCAAGTGGGCGAGATGATGAC 900  
Qy 1344 AACATGAGATGAGAGTCAACGTCGCCCTGACCGTCAAGCCCGGAGCGGCGGCG 1403  
Db 901 AACATGAGATGAGAGTCAACGTCGCCCTGACCGTCAAGCCCGGAGCGGCGGCG 960  
Qy 1404 GCGGAGCTCTGTCACAGGTCAGCGCGGCGCTCTCGGTGTCCTTTCAGAGAGCGC 1463  
Db 961 GCGGAGCTCTGTCACAGGTCAGCGCGGCGCTCTCGGTGTCCTTTCAGAGAGCGC 1020  
Qy 1464 GGGGGGGGTTGCGACCCCAAGAAAGGCTTGGCGGATCTTTTGGGCGGCTGCTG 1523  
Db 1021 GGGGGGGGTTGCGACCCCAAGAAAGGCTTGGCGGATCTTTTGGGCGGCTGCTG 1080  
Qy 1524 GGGGCTGTGGCCCTTACCGGTGTGGGAGAGTGAAGTGAAGACACCCGAGCGGCGC 1583  
Db 1081 GCGGCTGTGGCCCTTACCGGTGTGGGAGAGTGAAGTGAAGACACCCGAGCGGCGC 1140  
Qy 1584 CTGCTCTGCGCTCCCTCCCTGAGAAAAGACTCGGATGAGTGGGTCTGCGCTGT 1643  
Db 1141 CTGCTCTGCGCTCCCTCCCTGAGAAAAGACTCGGATGAGTGGGTCTGCGCTGT 1200  
Qy 1644 GCAAGGGAGTGGTCTTAAACCCTGTGTGATGAGTACAGCGCGTTTCAAGTCAC 1703  
Db 1201 GCAAGGGAGTGGTCTTAAACCCTGTGTGATGAGTACAGCGCGTTTCAAGTCAC 1260  
Qy 1704 ATGTGCTGGGCGAGACAGGTTTCTCTGTGCGCGGAGAAAGTTAACTTTCGCGC 1763  
Db 1261 ATGTGCTGGGCGAGACAGGTTTCTCTGTGCGCGGAGAAAGTTAACTTTCGCGC 1319  
Qy 1764 GGGCGTCAAGGCAATTAACGCTTACGAGAGGCTTATTCCTATTAATAGAAAC 1823  
Db 1320 GGGCGTCAAGGCAATTAACGCTTACGAGAGGCTTATTCCTATTAATAGAAAC 1379  
Qy 1824 GTCAAGTGAACCTTAATCTTCCGAGTTATGAGTTAACATGTGCTGTGGGCGTC 1883  
Db 1380 GTCAAGTGAACCTTAATCTTCCGAGTTATGAGTTAACATGTGCTGTGGGCGTC 1439  
Qy 1884 TTATACGGAGTCCGAGTCCGAGCCACCCCTGCGAGGTCGCGCCCTTTCGCGTGG 1943  
Db 1440 TTATACGGAGTCCGAGTCCGAGCCACCCCTGCGAGGTCGCGCCCTTTCGCGTGG 1499  
Qy 1944 ACAGTTTGAAGAGTGGGTGGAGTGAAGTTTGAAGGAGGAGCGTGTGTGTTCTA 2003  
Db 1500 ACAGTTTGAAGAGTGGGTGGAGTGAAGTTTGAAGGAGGAGCGTGTGTGTTCTA 1559  
Qy 2004 TGTGTTGGTCTGTTTCCCGAGCAAGAAATTCGATCAANTGTGACAGCTTTATTA 2063  
Db 1560 TGTGTTGGTCTGTTTCCCGAGCAAGAAATTCGATCAANTGTGACAGCTTTATTA 1619  
Qy 2064 CCTTAATCTTTCAGGGCTTAAATTTAGAGAGTGTCTGAGAGAGTTCATTAAGGGG 2123  
Db 1620 CCTTAATCTTTCAGGGCTTAAATTTAGAGAGTGTCTGAGAGAGTTCATTAAGGGG 1679  
Qy 2124 TTTCTTAAGACGCGTCAAGCCCTTCTAGAGAGTTATTCATTCGTCCCAAGAGCA 2183  
Db 1680 TTTCTTAAGACGCGTCAAGCCCTTCTAGAGAGTTATTCATTCGTCCCAAGAGCA 1739  
Qy 2184 GCTAGAGAGATTTGAGTCAATGACTTCCCACTGCGGCTCAAGGGGCTCAATTTAG 2243  
Db 1740 GCTAGAGAGATTTGAGTCAATGACTTCCCACTGCGGCTCAAGGGGCTCAATTTAG 1799  
Qy 2244 AAACCAAGAGGTTGAGTGAACCTTCAAGGCTTGAAGTTCAGTGTGCGACACTGC 2303  
Db 1800 AAACCAAGAGGTTGAGTGAACCTTCAAGGCTTGAAGTTCAGTGTGCGACACTGC 1859  
Qy 2304 CTGCGGAGAGGCTCTCCCAAGCAACCGAGATGGGGTTAAGAGAGAGAGAGCT 2363  
Db 1860 CTGCGGAGAGGCTCTCCCAAGCAACCGAGATGGGGTTAAGAGAGAGAGAGCT 1919

Qy 2364 TGGGTAGGGCCACTGTTTAAACAGCACTTCTCTCTGCGGCTTATTTTG 2423  
Db 1920 TGGGTAGGGCCACTGTTTAAACAGCACTTCTCTCTGCGGCTTATTTTG 1979  
Qy 2424 TTCAAGATGAGAGAGTGTGAACCTCTTTCAGAGAGGCTGGGAATCCTCTTAG 2483  
Db 1980 TTCAAGATGAGAGAGTGTGAACCTCTTTCAGAGAGGCTGGGAATCCTCTTAG 2039  
Qy 2484 AGCACTTAATCTTATTTATTCCTGAAATGTGCGTGTGCGCAGTGAAGAGGCTGGCTT 2543  
Db 2040 AGCACTTAATCTTATTTATTCCTGAAATGTGCGTGTGCGCAGTGAAGAGGCTGGCTT 2099  
Qy 2544 GCGAGCTCCCTAACCCTGCGCTGCGCCCTCCCGGGGTATGTGGCTTACCTGCGCA 2603  
Db 2100 GCGAGCTCCCTAACCCTGCGCTGCGCCCTCCCGGGGTATGTGGCTTACCTGCGCA 2159  
Qy 2604 CAGAGTTTGAAGCAATCAGCTGAGACTGAGTGGTTAAGATGTAACAGCTTAACTTGG 2663  
Db 2160 CAGAGTTTGAAGCAATCAGCTGAGACTGAGTGGTTAAGATGTAACAGCTTAACTTGG 2219  
Qy 2664 ATTTAAGAGCTTTTAAAGTATATCTCTGAAAGAAATGACGTAAACACAGGT 2723  
Db 2220 ATTTAAGAGCTTTTAAAGTATATCTCTGAAAGAAATGACGTAAACACAGGT 2279  
Qy 2724 GTACTATGAAAGCTGTATTTTAAATGAAGACGCTGGGCGAATGAATC 2771  
Db 2280 GTACTATGAAAGCTGTATTTTAAATGAAGACGCTGGGCGAATGAATC 2327

RESULT 7  
AK127646  
LOCUS AK127646 2327 bp mRNA linear PRI 19-FEB-2004  
DEFINITION Homo sapiens cDNA FLJ45744 f18, clone KIDNE2017153.  
ACCESSION AK127646  
VERSION AK127646.1 GI:34534649  
KEYWORDS oligo cloning; file (full insert sequence).  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.

REFERENCE  
AUTHORS  
1 Oshima, A., Takahashi-Fujii, A., Tanase, T., Imose, N., Takeuchi, K.,  
Arita, M., Mutsaers, K., Yuki, H., Hata, H., Sugiyama, T., Irie, R.,  
Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, T.,  
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,  
Matsumoto, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,  
Matsushima, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B.,  
Suzuki, Y., Sugano, S., Nagahara, K., Masuno, Y., Nagai, K. and  
Isogai, T.  
TITLE NEDO human cDNA sequencing project  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 2327)  
AUTHORS Isogai, T. and Yamamoto, J.  
TITLE Direct Submission  
JOURNAL Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7  
Unpublished

COMMENT  
Kazusa-Kametani, Kisei, Chiba 292-0818, Japan  
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)  
NEDO human cDNA sequencing project supported by Ministry of  
Economy, Trade and Industry of Japan; cDNA full insert sequencing:  
Research Association for Biotechnology (RAB); cDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: HRI and  
RAB; annotation: Reverse Proteomics Research Institute, HRI and  
RAB.

FEATURES  
source  
1. 2327  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"

/clone="KIDNE2017153"  
/cissue\_type="kidney"  
/clone\_lib="KIDNE2"  
/note="Cloning vector: pMR18SFL3"  
415..1122  
/note="unnamed protein product"  
/protein\_id="BAC87072.1"  
/db\_xref="GI:34534650"  
/translation="MARECKALLDGNKTTACTYHHVLTWGSSADSNQNRQLQKTR  
QKQBLAVSTCARLTAVLRDGLAADRAFERLWYAFSCLDLLEMDERALELGA  
PFLAPRPLRVRTGVAGASSGVARALSTSLRLAEAGDVEDVLELEFVQVEM  
IDNMEKVPRLPTVOARQAPAGAEELSTVSAGPSVSIQERGGCDPRALAAILFEG  
AVLAAVLAALVCAKLS"

CDS

ORIGIN

Query Match 73.9%; Score 2308; DB 8; Length 2327;  
Best Local Similarity 99.7%; Pred. No. 4,4e-290;  
Matches 2322; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

444 ACCGCGCGCGCTTGGAGGCGCTTAAGCGGAGCGCGGCTGTGCAAGCTGCTTGGCCC 503  
Db 1 ACCCGCGCGCGCTTGGAGGCGCTTAAGCGGAGCGCGGCTGTGCAAGCTGCTTGGCCC 60  
QY 504 GGAATTGGCAACCAAGAGATGGGACCGCACCTCAGCTTGCAGAGGACCAACGTCG 563  
Db 61 GGAATTGGCAACCAAGAGATGGGACCGCACCTCAGCTTGCAGAGGACCAACGTCG 120  
QY 564 AGGCCAGGCGGCTGACAGACACGACGTGTGACTCGAGTTCGCGCTTGGGAGAGATGACG 623  
Db 121 AGGCCAGGCGGCTGACAGACACGACGTGTGACTCGAGTTCGCGCTTGGGAGAGATGACG 180  
QY 624 AGGAGGCGGCGGACCGCTAAGCGGCGCTCTGCGCGCGCGCGCTCGCAGAGCGACGT 683  
Db 181 AGGAGGCGGCGGACCGCTAAGCGGCGCTCTGCGCGCGCGCGCTCGCAGAGCGACGT 240  
QY 684 CGAGGCTCCGCGCGGCGCTCGTGAACGTTGCGGTAAGCGCGAGCGAGTCAAGACAT 743  
Db 241 CGAGGCTCCGCGCGGCGCTCGTGAACGTTGCGGTAAGCGCGAGCGAGTCAAGACAT 300  
QY 744 GAAAGGCTTCTGTCG 803  
Db 301 GAAAGGCTTCTGTCG 360  
QY 804 CTGAGGCGGAGGCTTAAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 863  
Db 361 CTGAGGCGGAGGCTTAAGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420  
QY 864 AGGAGGAGTGAAGCGCGCTGTGACGAGGCTCAACAAGACGACTGCTGCTACACAC 923  
Db 421 AGGAGGAGTGAAGCGCGCTGTGACGAGGCTCAACAAGACGACTGCTGCTACACAC 480  
QY 924 CTGCTGTGAACCGTCCGCTGCTCGGCGGACTCGCAAGACCTGCGCGAGAGCTGCAAA 983  
Db 481 CTGCTGTGAACCGTCCGCTGCTCGGCGGACTCGCAAGACCTGCGCGAGAGCTGCAAA 540  
QY 984 ACCCGCGAAGAGCGCGAGAGCTGCGCGGTCCAACCTGCGCGCGCGCGCGCGCGCGCG 1043  
Db 541 ACCCGCGAAGAGCGCGAGAGCTGCGCGGTCCAACCTGCGCGCGCGCGCGCGCGCGCG 600  
QY 1044 CGGACCGGCGGCGCTGCGCGCGCGAGAGCGCGCGAGTTCAGAGCGGCTTGGTGGCTTC 1103  
Db 601 CGGACCGGCGGCGCTGCGCGCGCGAGAGCGCGCGAGTTCAGAGCGGCTTGGTGGCTTC 660  
QY 1104 TCGGCGCTGCTGGAACCTGTGGAAGCGGACATGCGACGCTGCGTGAAGCTGGCGCGCG 1163  
Db 661 TCGGCGCTGCTGGAACCTGTGGAAGCGGACATGCGACGCTGCGTGAAGCTGGCGCGCG 720  
QY 1164 TTCCCGCTGACGCGCGCGCGCGACGCTGCTGCGACAGATGAGCTGGCGCTGCTGCG 1223  
Db 721 TTCCCGCTGACGCGCGCGCGCGCGCGCTGCTGCGACAGATGAGCTGGCGCTGCTGCG 780  
QY 1224 GCGGTGCGCGCGCGCGCTGAGACCCGCGAGCTCGAGGCGGAGGCGGACTTC 1283

781 GCGGTGCGCGCGCGCGCTGAGACCCGCAAGCTCGAGCTCGAGCGGAGCGGACTTC 840  
QY 1284 GACGTGCGGACCTGTGGGAGAGCTGAGCGCGGAGTCTTCAAGTGGGCGAGATGAC 1343  
Db 841 GACGTGCGGACCTGTGGGAGAGCTGAGCGCGGAGTCTTCAAGTGGGCGAGATGAC 900  
QY 1344 AACATGAGATGAAGTCAACGTGCGCGCTGACCGCTGACCGCTGACCGCGAGCGCGCG 1403  
Db 901 AACATGAGATGAAGTCAACGTGCGCGCTGACCGCTGACCGCTGACCGCGAGCGCGCG 960  
QY 1404 GCGGAGCTCTGTCCAACGTCAAGCGCGCGCGCGCTCTGCTGCTGCTTTCAGAGAGCG 1463  
Db 961 GCGGAGCTCTGTCCAACGTCAAGCGCGCGCGCGCTCTGCTGCTGCTTTCAGAGAGCG 1020  
QY 1464 GGGGGGGGTTGGCAACCCCGAAGAGCGCTGCGCGCACTTTTCCGCGCGCGCTGCG 1523  
Db 1021 GGGGGGGGTTGGCAACCCCGAAGAGCGCTGCGCGCACTTTTCCGCGCGCGCTGCG 1080  
QY 1524 GCGGCTGTGGCCCTAGCGGTCGTGGCGAAGCTGAGCTGACAGACCCGAGCGCGC 1583  
Db 1081 GCGGCTGTGGCCCTAGCGGTCGTGGCGAAGCTGAGCTGAGACCCGAGCGCGC 1140  
QY 1584 CTGCTGCTGCGCTCCCTCCCTGAGAAAAGACTCGGATGGGTGTGGGCTTGGCTGT 1643  
Db 1141 CTGCTGCTGCGCTCCCTCCCTGAGAAAAGACTCGGATGGGTGTGGGCTTGGCTGT 1200  
QY 1644 GCAAGGGAGTGTCTTAACCCGCTGTGTGATGAGGTCACAGCCCGTTTCCAGTGCAC 1703  
Db 1201 GCAAGGGAGTGTCTTAACCCGCTGTGTGATGAGGTCACAGCCCGTTTCCAGTGCAC 1260  
QY 1704 ATCTGCTGGGACAGACAGGTTTCTCTGCTGAGCGCGGAGAGTTTACTTTCGCGC 1763  
Db 1261 ATCTGCTGGGACAGACAGGTTTCTCTGCTGAGCGCGGAGAGTTTACTTTCGCGC 1319  
QY 1764 GCGCGTCAAGGCAATTAACGCTTAACGCTGCAAGAGCTTTATTCCTATTATGAAGAAC 1823  
Db 1320 GCGCGTCAAGGCAATTAACGCTTAACGCTGCAAGAGCTTTATTCCTATTATGAAGAAC 1379  
QY 1824 GTCAAGTACCTTGAATCCCTCCGAGTTTAAGAGTTTACATATGCTGCTGGGCGCTC 1883  
Db 1380 GTCAAGTACCTTGAATCCCTCCGAGTTTAAGAGTTTACATATGCTGCTGGGCGCTC 1439  
QY 1884 TTTCAGGAGAGTCCGAGTTCGAGTCCCAACCGCTGCAAGGCGCGCGCTTTCGCGTGG 1943  
Db 1440 TTTCAGGAGAGTCCGAGTTCGAGTCCCAACCGCTGCAAGGCGCGCGCGCTTTCGCGTGG 1499  
QY 1944 ACAGTTTGAAGAGTGGGTGAGTGAAGTTTGAAGAGGACGCTGTGTTGTTCTA 2003  
Db 1500 ACAGTTTGAAGAGTGGGTGAGTGAAGTTTGAAGAGGACGCTGTGTTGTTCTA 1559  
QY 2004 TGTGTTGGTCTGTTTCCCGGACAGAAAAATTGCAATCAATGTCAAGAGCTTTATTA 2063  
Db 1560 TGTGTTGGTCTGTTTCCCGGACAGAAAAATTGCAATCAATGTCAAGAGCTTTATTA 1619  
QY 2064 CCTTAATCTTTCAGGCGCTAAATTTAGAGAGTGTCTGAGACGTTTCAATAAGAGGCG 2123  
Db 1620 CCTTAATCTTTCAGGCGCTAAATTTAGAGAGTGTCTGAGACGTTTCAATAAGAGGCG 1679  
QY 2124 TTTCTTGAAGCGCGCTTACGCGCTTCTAGCAGAGTTTATCATTTGTCCTCCCAAGACA 2183  
Db 1680 TTTCTTGAAGCGCGCTTACGCGCTTCTAGCAGAGTTTATCATTTGTCCTCCCAAGACA 1739  
QY 2184 GCTAAGAGATTTAGAGTCAAGACTCCCACTGCGCGCTCAAGGGGTGACCTTATTAAG 2243  
Db 1740 GCTAAGAGATTTAGAGTCAAGACTCCCACTGCGCGCTCAAGGGGTGACCTTATTAAG 1799  
QY 2244 AAACCAAGAGGATGAGTGAACCTTACTCTCAAGACTTGAATCCAGTGCACACTTTC 2303  
Db 1800 AAACCAAGAGGATGAGTGAACCTTACTCTCAAGACTTGAATCCAGTGCACACTTTC 1859  
QY 2304 CTGCGAAAAGGCTTCTCCCAAGCACCCGAGATGGGGTTAAGAGAAAGAGAGGCT 2363



Db 1860 CTGCGGAAAAAGGGCTCTCCACACACCCGAGATGGGGGTAGAGAGAGACAGAGCCT 1919  
Qy 2364 TGGGGTAGGGCCACTGGTGTATTAACAGGCACTTCTCTCTGGGGCTTATTTTG 2423  
Db 1920 TGGGGTAGGGCCACTGGTGTATTAACAGGCACTTCTCTCTGGGGCTTATTTTG 1979  
Qy 2424 TTGAGAACTAGACAGAGTGTGAACCTCTCTTGAAGAGGGCTGGGAATCCTCTTGA 2483  
Db 1980 TTGAGAACTAGACAGAGTGTGAACCTCTCTTGAAGAGGGCTGGGAATCCTCTTGA 2039  
Qy 2484 AGCACTTAATCTTATTTATCCCTGGAATGTGCTGTCAGTGAAGAGGGCTGGCTT 2543  
Db 2040 AGCACTTAATCTTATTTATCCCTGGAATGTGCTGTCAGTGAAGAGGGCTGGCTT 2099  
Qy 2544 GGCAGCTCCCTGACCCCGGCTGCGCCCTCTCCGGGGTAATGTGCAATTACTGGCCA 2603  
Db 2100 GGCAGCTCCCTGACCCCGGCTGCGCCCTCTCCGGGGTAATGTGCAATTACTGGCCA 2159  
Qy 2604 CAGAGGTTTGAACCAATGAGCTGAGACTGGGTTGAATGAAGAGCTTTAATCTTGGG 2663  
Db 2160 CAGAGGTTTGAACCAATGAGCTGAGACTGGGTTGAATGAAGAGCTTTAATCTTGGG 2219  
Qy 2664 ATTTAAGAGCTTTTAAAGATATATCTCTGAAGAAAGAAATGACGTAACCAAGCGT 2723  
Db 2220 ATTTAAGAGCTTTTAAAGATATATCTCTGAAGAAAGAAATGACGTAACCAAGCGT 2279  
Qy 2724 GTACTATGAAGCTGTATTTTATTAAGAAAGCTGGGCGATGAACTC 2771  
Db 2280 GTACTATGAAGCTGTATTTTATTAAGAAAGCTGGGCGATGAACTC 2327

RESULT 8  
AC027340 216441 bp DNA linear HTG IS-JUL-2000  
LOCUS  
DEFINITION Homo sapiens chromosome 19 clone CTD-2257C19, WORKING DRAFT  
AC027340  
SEQUENCE 52 unordered pieces.  
AC027340.2 GI:9211228  
HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
1 (bases 1 to 216441)  
DOE Joint Genome Institute.  
Sequencing of Human Chromosome 19  
Unpublished  
2 (bases 1 to 216441)  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (30-MAR-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
On Jul 15, 2000 this sequence version replaced gi:7341654.  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: <http://www.jgi.doe.gov>  
-----  
Project Information  
Center Project Name: 726304, BC691328  
Center clone name: CIRB-HI\_2257C19  
-----  
Summary Statistics  
Consensus quality: 164680 bases at least Q40  
Consensus quality: 187929 bases at least Q30  
Consensus quality: 195340 bases at least Q20  
Estimated insert size: 159060; agarose-fp estimation  
Estimated insert size: 211341; sum-of-contigs estimation  
Quality coverage: 8.59 in Q20 bases; agarose-fp estimation  
Quality coverage: 6.46 in Q20 bases; sum-of-contigs estimation.  
NOTE: This is a 'working draft' sequence. It currently  
\* consists of 52 contigs. The true order of the pieces

\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1 1041: contig of 1041 bp in length  
\* 1042 1141: gap of unknown length  
\* 1142 2270: contig of 1229 bp in length  
\* 2371 2470: gap of unknown length  
\* 2471 3552: contig of 1082 bp in length  
\* 3553 3652: gap of unknown length  
\* 3653 4683: contig of 1031 bp in length  
\* 4684 4783: gap of unknown length  
\* 4784 6178: gap of 1395 bp in length  
\* 6179 6278: gap of unknown length  
\* 6279 7442: contig of 1164 bp in length  
\* 7443 7542: gap of unknown length  
\* 7543 8661: contig of 1119 bp in length  
\* 8662 10030: gap of unknown length  
\* 8762 10030: contig of 1269 bp in length  
\* 10031 10130: gap of unknown length  
\* 10131 11491: contig of 1361 bp in length  
\* 11492 11591: gap of unknown length  
\* 11592 12967: contig of 1376 bp in length  
\* 12968 13067: gap of unknown length  
\* 13068 14173: contig of 1106 bp in length  
\* 14174 14273: gap of unknown length  
\* 14274 15840: contig of 1567 bp in length  
\* 15841 15940: gap of unknown length  
\* 15941 17242: contig of 1302 bp in length  
\* 17243 17342: gap of unknown length  
\* 17343 18449: contig of 1107 bp in length  
\* 18450 18549: gap of unknown length  
\* 18550 19683: contig of 1134 bp in length  
\* 19684 19783: gap of unknown length  
\* 19784 21247: contig of 1464 bp in length  
\* 21248 21347: gap of unknown length  
\* 21348 22952: contig of 1605 bp in length  
\* 22953 23052: gap of unknown length  
\* 23053 24452: contig of 1400 bp in length  
\* 24453 24552: gap of unknown length  
\* 24553 25825: contig of 1273 bp in length  
\* 25826 25925: gap of unknown length  
\* 25926 27109: contig of 1184 bp in length  
\* 27110 27209: gap of unknown length  
\* 27210 28292: contig of 1083 bp in length  
\* 28293 28392: gap of unknown length  
\* 28393 29744: contig of 1352 bp in length  
\* 29745 29844: gap of unknown length  
\* 29845 31042: contig of 1198 bp in length  
\* 31043 31142: gap of unknown length  
\* 31143 32629: contig of 1487 bp in length  
\* 32630 32729: gap of unknown length  
\* 32730 33898: contig of 1169 bp in length  
\* 33899 33998: gap of unknown length  
\* 33999 35704: contig of 1706 bp in length  
\* 35705 35804: gap of unknown length  
\* 35805 37471: contig of 1667 bp in length  
\* 37472 37571: gap of unknown length  
\* 37572 39285: contig of 1714 bp in length  
\* 39286 39385: gap of unknown length  
\* 39386 40577: contig of 1192 bp in length  
\* 40578 40677: gap of unknown length  
\* 40678 41930: contig of 1253 bp in length  
\* 41931 42030: gap of unknown length  
\* 42031 43743: contig of 1713 bp in length  
\* 43744 43843: gap of unknown length  
\* 43844 46054: contig of 2211 bp in length  
\* 46055 46154: gap of unknown length  
\* 46155 49379: contig of 3225 bp in length  
\* 49380 49479: gap of unknown length  
\* 49480 53371: contig of 3892 bp in length





Db 42804 GACCCCGCGCTCCGCCCTCCGCGGTAAATGTGGATTACTAGCCCAAGAGGTTTG 42863  
QY 2615 AGCCATCAGCTTGTGAGCTGGGTTTGATGATGACGCTTTAAGCTTTGAGTTTAAAGAC 2674  
Db 42864 AGCCATCAGCTTGTGAGCTGGGTTTGATGATGACGCTTTAAGCTTTGAGTTTAAAGAC 42923  
QY 2675 TTTTAAAGGTAAATTCCTGAAAGAAAATGAGTAAACGACGTTTACTATGAAA 2734  
Db 42924 TTTTAAAGGTAAATTCCTGAAAGAAAATGAGTAAACGACGTTTACTATGAAA 42983  
QY 2735 GCTGTATTTTAAAGAAAGCTGGGCAATGATCACTCCATGATGATCAAT 2794  
Db 42984 GCTGTATTTTAAAGAAAGCTGGGCAATGATCACTCCATGATGATCAAT 43042  
QY 2795 AGTATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2854  
Db 43043 AGTATCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 43102  
QY 2855 TAGTCTTCCCGCCCGCCCGCTTTTCTTTTGTGAGGAGTGAACCTGAGGCC 2912  
Db 43103 TAGTCTTCCCGCCCGCCCGCTTTTCTTTTGTGAGGAGTGAACCTGAGGCC 43160

RESULT 9  
AC151531 211799 bp DNA linear ROD 21-JUN-2005  
LOCUS Mus musculus BAC clone RP24-387D22 from chromosome 7, complete  
DEFINITION sequence.  
AC151531  
VERSION AC151531.2 GI:63004089  
KEYWORDS HTG.  
SOURCE Mus musculus (house mouse)  
ORGANISM Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 211799)  
Zheng, X. and Cotton, M.  
The sequence of Mus musculus BAC clone RP24-387D22  
Unpublished (2001)  
2 (bases 1 to 211799)  
Wilson, R.K.  
Direct Submision  
Submitted (10-SEP-2004) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
3 (bases 1 to 211799)  
Wilson, R.K.  
Direct Submision  
Submitted (09-MAR-2005) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
4 (bases 1 to 211799)  
Wilson, R.K.  
Direct Submision  
Submitted (03-MAY-2005) Genome Sequencing Center, 4444 Forest Park  
Parkway, St. Louis, MO 63108, USA  
5 (bases 1 to 211799)  
Wilson, R.K.  
Direct Submision  
Submitted (21-JUN-2005) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
On May 3, 2005 this sequence version replaced gi:51972398.  
COMMENT  
----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu>  
Contact: [submissions@wustl.edu](mailto:submissions@wustl.edu)  
----- Summary Statistics  
Center project name: M\_BB0387D22

## NOTICE:

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate  
chemistry, or covered by high quality data (i.e. phred quality  
>=30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone, fosmid clone or direct clone walk sequence.  
Sequence from the Mouse Genome Sequencing Consortium whole genome  
shotgun may have been used to obtain finishing sequence. The  
assembly was confirmed by restriction digest.  
This finishing standard has slightly changed from the previous  
human standard. Specifically, standards for regions of low sequence  
complexity (such as dinucleotide repeats and small unit tandem  
repeats) have been relaxed. These regions are very prevalent in the  
mouse genome, and the return on extended finishing efforts is  
minimal.  
If a sequence meets the criteria of the above statement, it needs  
no comments or tags. If the criteria are not met, such as ambiguous  
bases, then the region is duly annotated.

MAPPING INFORMATION:  
Mapping information for this clone was provided by Dr. Wes Warren,  
Department of Genetics, Washington University, St. Louis MO. For  
additional information about the map position of this sequence, see  
<http://genome.wustl.edu>

## SOURCE INFORMATION:

The BAC library has been constructed by Pieter de Jong and  
coworkers (<http://www.choi.org>) from male C57BL/6J mouse spleen  
and/or brain genomic DNA. The clone and detailed information can be  
obtained from Pieter de Jong and coworkers at <http://www.choi.org>

This sequence is the entire insert of the clone.

## FEATURES

source  
1..211799  
/organism="Mus musculus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10090"  
/chromosome="7"  
/clone="RP24-387D22"  
/clone\_1fb="RPCT-24"  
11573..11848  
/note="Unresolved tandem repeat."  
15154..15175  
/note="Sequence derived from one plasmid subclone."  
57981..57992  
/note="Sequence derived from one plasmid subclone."  
57987  
/note="Unresolved bases"  
100216..100326  
/note="Sequence derived from one plasmid subclone."  
100748..100942  
/note="Sequence derived from one plasmid subclone."  
118186..118246  
/note="Sequence derived from one plasmid subclone."  
203111..203365  
/note="Unresolved simple sequence repeat."

## ORIGIN

Query Match 24.1%; Score 753.8; DB 9; Length 211799;  
Best Local Similarity 64.5%; Pred. No. 5.7e-89;  
Matches 1597; Conservative 0; Mismatches 692; Indels 188; Gaps 24;  
QY 252 GTTTCCTCCATCGTAAATGAAAGGGTTGATCTCCGAGGCTTAATTCAGAACTC 311  
Db 59785 GTCGTGAAGTGAAGAAAAAACTTGAATTTGTAGGGCGCTTAAGGTTCAAGATTA 59844  
QY 312 GGATGGGGGGAAGGGAGGAGTGGGCCACCCACGTAACCTCCCGCGTGAAGCC 371  
Db 59845 TCAT-----GGTAGGGAAGAGAGGAAAGCCACACTTATCTACTGGGCTGGGCC 59895  
QY 372 CCGCTTACCACTGATTCAGGGGGTGGCAGCTCCGCGGAGACAGCGGGGTG----- 423  
Db 59896 CCGCTTACCTCGGCGCGGGGGGAGGGGGCGGGGGGCGGCGGAGAGAGAGCCCGG 59955





misc\_feature 217640..219158  
 /note="assembly\_name:Contig8"  
 gap 219159..219258  
 /estimated\_length=unknown  
 misc\_feature 219259..222531  
 /note="assembly\_name:Contig9"

ORIGIN

Query Match 24.1%; Score 753.8; DB 14; Length 222531;  
 Best Local Similarity 64.5%; Pred. No. 5,6e-89;  
 Matches 1597; Conservative 0; Mismatches 692; Indels 188; Gaps 24;

252 GTTTCCTCCATCCCTGTAATGAAACGGGTTGATCTCCGAGCGCTAACTTCCAGAACTC 311  
 Db 179225 GTCTGAAAGTGAAGAAAAAATACTTGAATTTGATGGCGCTTAAGGTTCCAGAAATTA 179284  
 312 GGAATGGGGCGAAGGGAGGAGGAGATGGCCACCCACACGTGAACCTCCCGCGCTGAGCC 371  
 Db 179285 TCAAT-----GGTAGGGAGAGAGAGGAAAGCCACACTTGAATCACTGAGGCTGGGCC 179335  
 372 CCGCCCTACGACTGATTCAGAGGGGTGGCACTCCGGCCGGAGCGAGCGGGGTG----- 423  
 Db 179336 CCGCCTAACCCCTGCGCCGGGGGAGGGGGGGGGGGGGGGGGGGGGAGAGAGGCCCG 179395  
 424 -----GGCGGGTCTAGGAAACCTTACCCGGCCGCTTGGCAGC 463  
 Db 179396 GGCTTGAGACTGGGTGGCTGAGAGGGCTGTCTTAGAAACTGAGATCTGTGCTTGGCCGG 179455  
 464 GCCTAAGGCGAGAGCGGCGGCTCTGAGAGCCTGCTTCCCGAGTTGGCAACCCACGAGG 523  
 Db 179456 GGCTAAGGCGGCGGCTCGCTTGA-----GACTTGGGGTGCATCAAG 179500  
 524 ATGGGAGCGGACCCCTCAGCTTTCAGAGGAGCCACTGTGAGGGCCAGGGCGGTGCAAGA 583  
 Db 179501 ATGGGGTCTTCACTCTGGGCTTCTCAGGAGCGCGGTGAAAGAAATGTGTGCAAGA 179560  
 584 CACGAGTGTGACTCGGAGTGGCGCTGGGAGAGATGAGACGAGGAGCGGGGAGCCGCTAA 643  
 Db 179561 CACTAGCGCTGACCCGAGAGCGCTCTGGGGGAATGAGAGCGGATGGGAGAGACTTA 179620  
 644 CGGGGCTCCCTTGGCGGCGCCGCTCGAGAGGCGGAGCTGAGAGGTCCTGGCGGGCTC 703  
 Db 179621 CCAAGCTACCCAGC-CGTGGGAGACCGCGAGGCGCACTCAAGGGTCCCTGAGAGCTC 179679  
 704 CGTGAGACTTGGCGGTGAGCGCGAGGAGTCAAGACATGAAAGAGCTTTCGTCGGCCG 763  
 Db 179680 TGTGTGAACAGCGG-----CGAGGACATGAGAGCTTTCAGTTCGCC 179723  
 764 GGCCTAAGCGGAGATGGGGGTAGCCATCTCTGCGCGCTGAGGGGAGGCTAACGGG 823  
 Db 179724 TGTCTAAGGCTGCACTCAGGGCTGAGAGGCACTTACGAGCTTGAG----- 179769  
 824 CGCGGCGGCGCGGAGCCAGCGGAGCCACCGGATGGGAGGAGAGATGCAAGCGCT 883  
 Db 179770 -GTTGCAATGAGGTTACCGCGGAGGCAAGCGCATGCGCAAGAGAGTGCAGAGCGCT 179828  
 884 GCTGAGCGGCTTCAAGAGCACTGCTGTACCAACCTGTGTCTGACCGTCCGTGG 943  
 Db 179829 GCTGAGCGGCTTCAAGAGCACTCAGCGGTCTACCTTACCTTATGTGTGGCGG 179888  
 944 CTGCGGAGCTCGCAGAGCTGCGGAGAGCTGCAAAAGACGCGCGAGAGGCGCAGGA 1003  
 Db 179889 CTCCGCGAGCACTCAGAGCTTACGCGAAGCTTGCAGAAAGCCCGCAGAGAGCGGGGA 179948  
 1004 GCTGGCGGTGTCACCTTGGCGCGGCTGACTGTGTGTGTCGCGAGCGGGGCGCTGGCCG 1063  
 Db 179949 GCTGGCGGTGTCACCGGCGCGGCTTGAACGTCGCGCTCCGGAGCCGAGGCTTACCTAC 180008  
 1064 CGACGAGCGCGCGAGATTTGAGAGCGCTTGGGTGGCTTTCAGGGTGTGAGACTTGT 1123  
 Db 180009 CGAGGAGCGCGCGAGATTTGAGAGCGCTTGGGTGGCTTTCAGGGTGTGAGACTTGT 180068  
 1124 GGAAGCGGAGCATGCGAGCTGCGTGAAGCTGGGCGCGGCTTCCGCTCAGCGCGCGG 1183

180069 CGAAGCGGAGCATGCGAGCGGCACTGAGCGCTGGGAGGCACTTCCCGCTGACAGCGCGCG 180128  
 1184 GCGACCGCTGTGTGCGCAGAGGTGTGGCTGTCCCTCTCCGCGTGTGGCGCGCGCT 1243  
 Db 180129 CCGGCGCACTGTGTCCACTGTGGGTGTAAGCGGCGGCTCTCTGTGTGATGCTGCGCGCGCT 180168  
 1244 GAGCAGCGGAGCTGTGCGGCTGAGGCGGAGGCGGAGCTTTCGACGTGCGGAGCTTGGCGG 1303  
 Db 180189 GAGCGCGGAGCTGTGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180248  
 1304 GCTGAGCGCGAGGCTCTTCAAGGTGGCGGAGTGTGCAAAATGAGATGAGAGTCA 1363  
 Db 180249 GTTGGAGCGCGAAGTCTCAGGTGGGCGGAGTGTGAGCAGCATGGAATGAAATGCA 180308  
 1364 CGTCCCGCGCTGAGCCGTGCAAGCCCGGAGCGCGCGCGCGCGAGCTCTCTTCCACGGT 1423  
 Db 180309 CGTCCCGCGCTGAGCTGTACAGGCGGAGCAGAGCGGAGCGCGGAGCTCTCTGTCGGTGC 180368  
 1424 CAGCGCGGCGCGCTCTCGGTG-----TGTCTTGCAGAGAGCGCGGGGGGTTGCCA 1477  
 Db 180369 CAGTCCGCGCGCTCTCAGCGGTGCAATTCGTTAGAGAGCGCGCGCGCGCTTCCAG 180428  
 1478 CCCAGAGAGCGCTGAGCGCGCATCTTTCGAGCGCGCTGCTGCGCGCTGTCGCGCT 1537  
 Db 180429 CCCAGAGAGCGCGCTGAGCGCGCATCTTTCGAGCGCGCTGCTGCGCGCTGTCGCGCT 180488  
 1538 AGCGGTGTGCTGGCGAGAGTGAAGCTGACAGACCCGAGCGCGCTGCTGCGCGCT 1597  
 Db 180489 CGCTTATGTGTGGCAAGCTGAGCTGATGATCAACGAGTGTGCGCGCTGCTGCTT 180548  
 1598 CCTCCCTCGAAGAAAGCTCGGAGTGGGTGTGGGTGTGGGTGTGGGTGTGGGTGTGGGTGTGG 1657  
 Db 180549 CTTTC-----CAGGCAATGGGCTTGGGTGTGAAACCG-TAAGAGAGAGGT 180594  
 1658 CCTAAACCCCGTGTGTGATGAGTACAGCGCGCTTC-CAGTGCACATCTGCTGGGAG 1716  
 Db 180595 AGTGAAGCGTGAATGTGTGATGAGATGAGATGAGATGAGATGAGATGAGATGAGATGAG 180654  
 1717 GAGC-----ACGTTTTCTCTTGTGTGCGCGCGGAGAGAGTGAATCTTGGCGCGCGCTCA 1771  
 180655 AAGCGT 180713  
 1772 GGGCAATACCGCTAAGTGTGAGAGAGCTTATTCCTATTATGAAACCGT-CACAG 1830  
 Db 180714 GGCATTTACTTATATGT 180773  
 1831 TGACCTTGAATCCTCCGAGTTAATGATTAACATGTGCTGTGGGCGCTCTTAAAG 1890  
 Db 180774 TGACCCAGGTCCCTGAGT-----TTGATGGGTTAAGAGCT 180821  
 1891 GAGATCCAGATTCGATGCCACCCCTGCGAGCGTGCCTCTTGTGCTGGAGCAGTTT 1950  
 Db 180812 ATGTCCCAAGTTGTGCTTACCGCTGAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 180871  
 1951 GAAAGT 2010  
 Db 180872 AATTTGT 180926  
 2011 GGT 2068  
 Db 180927 ATTTGT 180986  
 2069 ATCTTTCAGGGCTTA-AATTTAGAGAGTGTCTTGAAGAGCTTCAATCAAGAGGCTTTC 2127  
 Db 180987 ATCTTTCAGGGCTTGT 181046  
 2128 TCTAAGAGCGCTAAGCGCTTCTAAGAGAGTGTATCAATTCGTCCCAAGAGAGCTA 2187  
 Db 181047 TCTGTATCACTTACTGACCTTG---GCAAGCTGCTTATTAATCAAGAGAGCTG--- 181101  
 2188 GAAAGATTTTGAAGTCAATGACTTCCCACTGCGCTCAAGGAGCTGACCTTATTTAGAAAC 2247

Db 181102 -----AGCATGAGAGGTTTTTGTGTGTCGAGCCCATTTAGGGTCA 181142  
QY 2248 CAAAGAGGTTGGTTGACCTACTCTCAGCAGACTTGATCAGTGGCACAATTGGCTTGC 2307  
Db 181143 CAGCCAGAGAGGTTGTGTGTATCTC-----TGCACTTGTCTTCA 181184  
QY 2308 GGAAGAGGCTCTCCCGCAGCCCGGAGATGGGGGTAAAGAGAGAGCAGAGGCTTGG 2367  
Db 181185 AGAGAGAGCCGCTCCACTACTCTCCCGATGGGGAGAGAGAGTGTCTGCTTGG 181244  
QY 2368 GTAGGGCCACTGTG-TGTTTAAACAGAGCACTTCTCTCTGAGGGCTTAATTTTGTTC 2426  
Db 181245 GAAAGGCCACACGGCTATATGAAACAGCAAGTCTCTCTCTCAGCGGCTTAATTTCC 181304  
QY 2427 AGAATGACACAGAGTGTGAACTCTCTTGTGAGAGAGGCT-GGAAATCTCTTTAGAG 2485  
Db 181305 AGAATGACACAGAGTGTGAACTCTCTTGTGAGAGAGGCTGAGGAATCTCTTTGAG 181364  
QY 2486 CACTTAATCTTATTTATCCCGGAATGTGCGTGTGCGCAGATGAGAGGCTGAGCTTGG 2545  
Db 181365 CTCTTAATCTTATCTAT-CCCTGGAATTTACTGATAGCCATAGAGAGCTAG-TTTTC 181422  
QY 2546 CAGCTCCCTGACCCCGCGCTGCGCCCTCCGGGTAATGTGCAATTACTGGCCACA 2605  
Db 181423 TAGAGAGCTGCCCACTGTCTGACCCCATGTTACAGAGAGGCGAGTGTATCACTA 181482  
QY 2606 GAGGTTTGAAGCAATCAGCTCTGAGACCTGGGTTAAATGTAAACAGCTTAACTTGGAT 2665  
Db 181483 GTAAACGAGGGGTTGTGAGGCTGAGCACTGAGTTCAAAATGTAAACAGCTTAAATTTGAAAT 181542  
QY 2666 TTAAAGAGCTTTTAAA 2682  
Db 181543 TTTAAAGCTTTTAAACA 181559

RESULT 11  
AY229888 6560 bp mRNA linear ROD 01-JAN-2004  
LOCUS AY229888  
DEFINITION Mus musculus RGS9-1 anchoring protein R9AP mRNA, complete cds.  
ACCESSION AY229888  
VERSION AY229888.1 GI:37910119  
KEYWORDS  
SOURCE Mus musculus (house mouse)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
Sciurognathi; Muridae; Murinae; Mus.  
REFERENCE  
AUTHORS Kereztes, G., Mutai, H., Hibino, H., Hudspeth, A.J. and Heller, S.  
TITLE Expression patterns of the RGS9-1 anchoring protein R9AP in the  
chicken and mouse suggest multiple roles in the nervous system  
Mol. Cell. Neurosci. 24 (3), 687-695 (2003)  
JOURNAL  
PUBMED 14664818  
2 (bases 1 to 6560)  
Kereztes, G. and Heller, S.  
AUTHORS Direct Submission  
TITLE Submitted (03-FEB-2003) Otolaryngology, Harvard Medical School, 243  
JOURNAL Charles Street, Boston, MA 02114, USA  
FEATURES  
source  
1. 6560  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL"  
/db\_xref="taxon:10090"  
/feature\_type="retina"  
325..1038  
/note="membrane-spanning; SNARE-like"  
/codon\_start=1  
/product="RGS9-1 anchoring protein R9AP"  
/protein\_id="AAP5847.1"  
/db\_xref="GI:37910120"  
/translation="MARCKSKALDALNTKTAACYHILVITVGSGADTODLRBELQKTR  
OKARELAIVATGARLTVALRDRSLATERBAEFERLVAVAFSGCDLLEADMQRALALCAT

## ORIGIN

FFLHARRPLVITVGSGSAVAPALASRSIRHREBSDFVADLIPOLEREVLQVEM  
IDDEMKYVNPRTVQARQAAAGALLSASASASAGSISVERBAGPDPSALAAATV  
FSAVLIVVALALCVAKLS"

Query Match 23.6%; Score 738; DB 9; Length 6560;  
Best Local Similarity 66.0%; Pred. No. 1.3e-86;  
Matches 1450; Conservative 0; Mismatches 610; Indels 136; Gaps 21;

QY 505 GAGTTGACCCACGAGAGATGGGACCGACCTCTGAGTTGCGAGAGGACCACTGTGA 564  
Db 4 GACTTGGGGTCGATCAAGATGAGGCTTCACTCTGAGCTTCTCAGAGGCGCGGTGAA 63  
QY 565 GGCACGGCGGTGACAGACACAGAGTGTGATCGAGAGTGGCGCTGGAGAGAGATGGAGCA 624  
Db 64 AAGAAAGTGTGTCAAAAGCACTACCGGTGACCGGAGGCGCTCTGGGGAAATGAGAGA 123  
QY 625 GGAAGCGGGAGACCCCTAAAGGGGCTCCTCTGCGCGCCCGTCCGACAGAGCGACGTC 684  
Db 124 GCGAATGGGAGAAAGCTAACAGACTAACCCAGC-CGTGGGAGACCGCGAGCGCACCTC 182  
QY 685 GAGGTTCCCGGCGGGCTTCGTGACGTTGGCGGTAGCGCCGAGAGTCAAGACATAG 744  
Db 183 AAGGTTCCCTGGAAGAGCTCTGTTGAAACAGCGG-----CGAGGACATAG 226  
QY 745 AAGAGGTTGCTGTCGCGCGCGCCCAAGGCGGGATGGGGTTTAAAGCAATCTGCGCGAG 804  
Db 227 AGAGGTTTCAATCGGCTCTGCTCAAGGCTGTGACATCAAGGCTGTGAGGCGCACTTAAGAGT 286  
QY 805 TGAGGGGAGAGGCTTAAAGGGCGCGGCGCGGCGCCAGCGGAGCCACCGGATAGCGCA 864  
Db 287 TGAG-----GTTGCAAGTAGGGTACCGCGGAGGCGAGCGCATAGCGCA 331  
QY 865 GGAAGAGTGTCAAGGCGCTGTGTGACAGGCTCAACAAAGCACTGCGTCTACCAACACC 924  
Db 333 GAGAGGAGTGCAGAGCGCTGTGTGACAGGCTCAACAAAGCACTGCGTCTACCAACACC 391  
QY 925 TGGTGTGACGCTGCGGTGTGCTGCGGAGCTGCGAGAACCTGCGGCGAGAGTGCAGAAAGA 984  
Db 392 TAGGTGTTACTGTGGCGGCTCGCGGACACTCAAGACCTTACGCAAGAGCTGTGAGAAAGA 451  
QY 985 CCGCGCAGAAAGCGAGAGAGTGTGTGTTCACCTGTGCGCCCGGCTGACTGTGTGTGTC 1044  
Db 452 CCGCGCAGAAAGCGGAGAGCTGTGTGTTCACCGCGCGCCGGTGTGACGCTGTGCGCTTC 511  
QY 1045 GCGACCGGGGCTGTGCGCGCCACAGAGCGCGCGAGTTGAGAGGCTCTGTGGTGGCTTCT 1104  
Db 512 GCGACCGGAGCTTACTTACCTACCGAGAGCGCGCGAGATTGAGCGGCTGTGGTGGCTTTT 571  
QY 1105 CGGGCTGCTGTGACCTGTGAAAGCGGACATGCGAGCTCGCTGTGAGCTGTGGCGCGCGT 1164  
Db 572 CAGGCTGTGTAAGCTGTGTGCAAGCGGCGCGAGATTGAGCGGCTGTGGTGGCTTTT 631  
QY 1165 TCCCGTGTGACCGCGCGCGCGAGACCGCTGTGTGTGCGACAGTGTGTGTGTGTGTGTG 1224  
Db 632 TCCCGTGTGACCGCGCGCGCGCGAGCACTTGTGTGACTGTGTGTGACGCGGCTGTGTG 691  
QY 1225 GCGTGGCGGCGCGCGCGCGCTGAGCACCCGAGAGCTGTGTGTGTGAGCGGAGCGACTTG 1284  
Db 692 CTGTAGCTGTGCGCGCGCTGAGCGCGCGAGCTGTGTGTGTGAGCGGAGCGAGACTTGT 751  
QY 1285 AGCTGCGGACCTGTGCGGAGCTGTGAGCGCGAGGTCTCTTCAAGTGTGGCGAGATGATGACA 1344  
Db 752 ATGTGCGCATGTGCGCGCGCGGTTGAGAGCGGAAGTCTTCAAGTGTGGCGAGATGATGAG 811  
QY 1345 ACATGAGATGAAAGTCAACGTGTGCGCGCGGTGACCGGTGACCGCGCGCGCGCGCG 1404  
Db 812 ACATGAGATGAAAGTCAACGTGTGCGCGCGGTGACCGGTGACCGCGCGCGCGCGCG 871  
QY 1405 CCGAGCTCTGTGTCAAGGTCAAGCGCGCGCGCGCTCTCTGTGTG-----TGTCTGTGAGG 1458  
Db 872 CCGAATCTCTGTGTGCGGTGCAAGTGTGCGCGCGCGCTCTCAAGCGGTGCAATTTGCTGTAGAG 931



QY 1459 AGCGGCGGGGGGTGCGAAGCCCGAGAGAGCCCTGCGCCGATCTTTTGCGCGCGTGC 1518  
 DB 932 AGGCGCGGGGCGCCCTGCGAAGCCCGAGAGAGCCCTGCGCGAAGCTTTTCAAGCGCGTGC 991  
 QY 1519 TGTGCGCGGTGTGGGCTTACGCGGTGTGCGTGGCGAAGCTGACGACAGACCCGACG 1578  
 DB 992 TGTGCGGTGTGGGCTTACGCGGTGTGCGTGGCGAAGCTGACGACGACG 1051  
 QY 1579 GCGCGCTGCTGCGCGCTCCCTCCCGTGAAGAAAGCTGGAGTGGGTGTGGGCTGCG 1638  
 DB 1052 GTGCTCGCGTCCCTTTTCTTTC-----CAGGCAATGGGCTTCGGGTTTGA 1098  
 QY 1639 CCTGTGCAAGGGAGTGTCTTAAACCCCGTGTGCAATGGAGTGAACGCGCTTTC-CA 1697  
 DB 1099 ACCGTT-AGGAGAGAGGTAGTGAAGCTGAATGTGCAATGAATGATCTTTCAAA 1157  
 QY 1698 GTGCAATCTGCTGTGGCGAGAC-----ACGTTTCTTGTGCTGCGCGGAGAAATT 1752  
 DB 1158 GCCCACTCTGCTGTGGCAAAAGTGTGTTTAACTGCTGCGCCCAAGAGAGATT 1217  
 QY 1753 AACTTGGCGCGGCGGTGAGGGCATTAACGCTAAGCTGAGAGCTTTATTCCTATT 1812  
 DB 1218 AACAGAC-AGGCGCGCGAGGCAATTAATCTTAATGTGTAGAGCTTTAATCTCTATT 1276  
 QY 1813 AATGAGAAACCGT-CAAGTGAACCTAGATCCCTCCGAGTTAATGAGTTAAACATGTGC 1871  
 DB 1277 AATGAGAAATCGGTCCAGTGAAGCCAGGCTCCTGAGT----- 1316  
 QY 1872 TGTGGGCGGTCTTTACAGGAGTCCGAGTCCGTCGCAACCCCTGCGACGCTGCGCC 1931  
 DB 1317 --TTGATGGGTAAAGATGATGTCCCAAGTTGTGTCCTCCAGCGTGTGTGAGCACTC 1374  
 QY 1932 TTTTCGCTGGGACAGTTTGAAGAGGAGTGGGTGGAGTGAAGTTTGGAGAGGACGC 1991  
 DB 1375 ATTCTTTTAAAGCAGACAAATTTTGAAGGTGTGGGTGTATTAAGATATGTGTG 1434  
 QY 1992 TGTGTTGTTCTATGTGTGTGTCTGTTCCCGAGCAAGAA--AATGCAATCAATGTCT 2049  
 DB 1435 TGGGGG-----AGGAGACATTTTGTTCCTTAAGCAAGAAACAGTTGAGTAAATGG 1489  
 QY 2050 AGCAGTTTATTAATCTTAATCTTTACAGGCTTA--AATTTAGAGAGTGTCTGAGACA 2108  
 DB 1490 CGTGGGTTTGTATCACTTAATCTTTGCGGCGTGAAGCTTTGGAGTGTGATGATGAG 1549  
 QY 2109 GTTCAATACAAAGGCTTCTCTAAGACGGCTACAGCCCTCTTACACAGTTTATTCAT 2168  
 DB 1550 GCTCGTGAAGGCTTCTCTGTGTATCACTAATCACTAATCTTGTGCTTAT 1606  
 QY 2169 TGTGCTCCCAAGAGCAGTGAAGAGATTGAGGTGATGACCTCCACTGCGCTCAGGG 2228  
 DB 1607 TATGTACCAAGAGCTGC-----AGCTATGAGAGTTTGTGTCTC 1645  
 QY 2229 CTGACCTTATTTAGAAACCAAGAGGTTGAACTTCTGACGAGCTTGTGATCC 2288  
 DB 1646 GTGACCCCAATTTAGGCTCAAGCAAGAGGTTGTGTATCTCT----- 1692  
 QY 2289 AGTGGCAACTGCTGTGGGAAAGGGCTCTCCAGCCACCCGAGAGTGGGGTAA 2348  
 DB 1693 ----GCAACTTGTCAAGAGAAACCGCTCATATCTATCCCGAGTGGGAGGGA 1747  
 QY 2349 GGAAGAGCAGAGCTTGGGTTAGGGCAC-CTGTGTTTAAACAGGACATTTCTCTCT 2407  
 DB 1748 CAGGATGTGTCTTGTGGGAGGGGCAACCGCTATGTAAACAGGACGTTCTCTTCT 1807  
 QY 2408 CTGGGCTTATTTTGTCTTGAACCTAGACAGTGTGAACTCTTTTGCAGAGGGC 2467  
 DB 1808 CACCGCTTACTTATCCAGATTTGACACAGTGTGAACTCTCTTTGAGGGAGGCG 1867  
 QY 2468 T-GGGAATCTCTTTAGAGCACTTAATCTAATTTATCCCTGGAATGTGTGCTGGCA 2526  
 DB 1868 TGGGAATCTCTTTAGAGCTTAACTTAATCTAAT-CCCTGGAATTTACTGATAGCCA 1926  
 QY 2527 GTAGAGAGGCTGTGCTTTGAGAGCTTCCCTGACCCCGCGCTGCCCGCTCTCGGGGTAA 2586

DB 1927 ATGGAACCTAG-TTTTCTAAGAGAGCTCCCAATGTCTGAGCCCATGTTACAGGAG 1985  
 QY 2587 GTGCACTTACTGCGCCACAGAGGTTTGAAGCCCAATAGCTGTGAGCTGGTTAGATGT 2646  
 DB 1986 GCGGATGCTGTATCACTAGTAACAGGGGTTGTGAGCCCTGCGAGCTGAGTTCAAAATGT 2045  
 QY 2647 AACAGCTTAACTTGGGATTTAAGAGCTTTTAAA 2682  
 DB 2046 AACAGCTTAACTTGGGATTTAAGAGCTTTAACA 2081  
 RESULT 12  
 AC128498/c  
 LOCUS  
 DEFINITION  
 Rattus norvegicus clone CH230-328014, \*\*\* SEQUENCING IN PROGRESS  
 \*\*\* 3 unordered pieces.  
 AC128498  
 AC128498.3 GI:25138075  
 HTG; HTGS PHASE1; HTGS DRAFT; HTGS\_ENRICHED.  
 Rattus norvegicus (Norway rat)  
 ORGANISM  
 Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;  
 Sciurognathi; Murioidea; Muridae; Murinae; Rattus.  
 1 (bases 1 to 255844)  
 Muzny, D., Marie, Metzker, M., Lee, Abramson, S., Adams, C., Alder, J.,  
 Allen, C., Allen, H., Albrooks, S., Amth, A., Anguitano, D.,  
 Anyalbech, V., Aoyagi, A., Ayodeji, M., Baca, E., Baden, H.,  
 Baldwin, D., Bandaranaike, D., Barber, M., Barnstead, M., Benham, F.,  
 Bernal, K., Blair, J., Blankenburg, K., Blyth, P., Brown, M.,  
 Bryant, N., Buhay, C., Burch, P., Burrell, K., Calderon, B.,  
 Cardenas, V., Carter, K., Cavazos, I., Caesar, H., Center, A.,  
 Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Y., Chen, Z., Chu, J.,  
 Cleveland, C., Cockrell, R., Cox, C., Coyle, M., Cree, A., D'Souza, L.,  
 Davila, M. L., Davis, C., Davy-Carroll, L., De Anda, C., Dederich, D.,  
 Delgado, O., Denson, S., Deramo, C., Ding, Y., Dinh, H., Dlyva, K.,  
 Draper, H., Dugan-Rocha, S., Dunn, A., Durbin, K., Duval, B., Eaves, K.,  
 Egan, A., Escoto, M., Eugene, C., Evans, C. A., Falls, T., Fan, G.,  
 Fernandez, S., Finley, M., Flagg, N., Forbes, L., Foster, M., Foster, P.,  
 Frazer, C. M., Gabisi, A., Ganta, R., Garcia, A., Garner, T., Garza, M.,  
 Gebregorgis, E., Geer, K., Gill, R., Grady, M., Guerra, W., Guevara, W.,  
 Gunaratne, P., Haaland, W., Hamill, C., Hamilton, C., Hamilton, K.,  
 Harvey, Y., Havlak, P., Hawes, A., Henderson, N., Hernandez, J.,  
 Hernandez, R., Hines, S., Hladun, S. L., Hodgson, A., Hogue, M.,  
 Hollins, B., Howells, S., Hulik, S., Hume, J., Idlebird, D., Jackson, A.,  
 Jackson, L., Jacob, L., Jiang, H., Johnson, B., Johnson, R., Jolivet, A.,  
 Karpathy, S., Kelly, S., Kelly, S., Khan, J., King, L., Kovar, C.,  
 Kowis, C., Kraft, C. L., Lebow, H., Levan, J., Lewis, L., Li, Z., Liu, J.,  
 Liu, J., Liu, W., Liu, Y., London, P., Longacre, S., Lopez, J.,  
 Lorenshew, L., Louised, H., Lozano, R. J., Lu, X., Ma, J.,  
 Maheshwari, M., Maindarte, M., Mahmoud, M., Malloy, K., Mangum, A.,  
 Mangum, B., Mapua, P., Martin, K., Martin, R., Martinez, E.,  
 Manning, S., McLeod, M. P., McNeill, T. Z., Meenen, E.,  
 Milosavljevic, A., Miner, G., Mirza, E., Montemayor, J., Moore, S.,  
 Morgan, M., Morris, K., Morris, S., Mundiada, M., Murphy, M., Nair, L.,  
 Nankervis, C., Neal, D., Newton, N., Nguyen, N., Norris, S.,  
 Nwankwelen, O., Okunodu, G., Olarnungsoom, A., Pal, S., Parks, K.,  
 Pasternak, S., Paul, H., Perez, A., Perez, L., Pfannkuch, C.,  
 Plopper, F., Poindexter, A., Popovic, D., Primus, E., Pu, L.,  
 Puazo, M., Quiroz, J., Rachlin, B., Reeves, K., Regier, M. A., Relph, R.,  
 Reilly, B., Reilly, M., Ren, Y., Reuter, M., Richards, S., Riggs, F.,  
 Rives, C., Rodkey, T., Rojas, A., Rose, M., Rose, R., Ruiz, S. J.,  
 Sanders, M., Savery, G., Scherer, S., Scott, G., Shatman, S., Shen, H.,  
 Shetty, J., Shvartsbeyn, A., Sison, I., Sitter, C. D., Smjs, D.,  
 Sneed, A., Sodergren, E., Song, X.-Z., Sorelle, R., Soza, J.,  
 Steimle, M., Strong, R., Sutton, A., Svatek, A., Taboz, P., Taylor, C.,  
 Taylor, T., Thomas, N., Thomas, S., Tingey, A., Trejos, Z., Umami, K.,  
 Valdes, R., Vera, V., Villasana, D., Waltron, L., Walker, B., Wang, J.,  
 Wang, Q., Wang, S., Warren, J., Warren, R., Wei, X., White, F.,  
 Williams, G., Wilson, R., Wiczysk, R., Woodem, H., Wolley, K.,  
 Wright, D., Wright, R., Wu, J., Yakub, S., Yen, J., Yoon, L., Yoon, V.,  
 Yu, F., Zhang, J., Zhou, J., Zhou, X., Zhao, S., Zhao, S., Dunn, D., von



[illegible]

QY	1996	GGGTTCTAATGNGGTTGCTGTGTTCCCGCAAGAAGAAATTCGAATCAATGTCACAGC	2055
Db	173027	TGTTCCCTAG-----ACAGAAAGCAACTGTAGATGTGC	172990
QY	2056	TTTTATTAACCTTAATCTTTCAGGGCGCTAAATTTAGAGAGTGTCTTGAGAGCACTTCATA	2115
Db	172989	GTGGGTACTTAAATCTTCCTGGGGCCTTAAGCTTTTGAGAGAGTGTATCTGAGAGAGTTTCATG	172930
QY	2116	CAAGGGCTTTCTCTTAAGACGCGCTACAGCCCTTCCTAGAGATTATTCATTGCTCCC	2175
Db	172929	CAAGGTGCTTCTCTGTGGGTACACTATTTGCGCTGTGGCAGAGATTCGCTTATTTGTCACAG	172870
QY	2176	CAAGGACGCTAGAGAGATTTGAGAGTATGACCTCCACCTGCGCTGAGGGGCTACCC	2235
Db	172869	GAACTGCAGCTCTGAGAGGCTTTGTGTGTGAC-----CATTATGGGTCACAAACC	172818
QY	2236	TATTTAGAAAACCAAGAGGGGTGGGTTGAACCTACTCTCACGGAATTGGATCCAGTGCGC	2295
Db	172817	AAAAAGTTGCTTCAATTCGATGTGTGTATTCACCT-----GC	172780
QY	2296	ACACTTGCCTGCGAAAAAGGCTCTCCCGACCCGAGATGGGGGTAAAGAGAAAG	2355
Db	172779	ACACTGCGCGCAATTAGAAACGCGCTCCCTACCCATCCCTGATGGGGGAGAAACAGATG	172720
QY	2356	CAGAGGCTTTGGGGTGGG--CCACTGTGTGTTAAACAGGCACTTCTCTCTCTGGGGC	2414
Db	172719	CGCGCTTTTGGGAAAGGGCCCTCACTGAGATGTAAACAGTCAAGTCTTCTTCTCACCGGC	172660
QY	2415	TTATTTTGTTCAGAACTAGACAGAGTGTGAACCTCCTTGTGAGAGGGGT--GGAA	2473
Db	172659	TTACTTTATCCAGAAATTAGACACAGTGTTCGACCTCTTGTGTGGAGGGCTGGGGAA	172600
QY	2474	TCTCTTTTAAAGCACTTAATCTTATTTATCCCTGGAATGTGCGTGTGGCCAGTAGAG	2533
Db	172599	TCTCTTTGAAGCTCTTAATCTTAATCTTAATCCCG--GGAATTACTCAATAGCCAAATGGAG	172541
QY	2534	GAGTGGCTTTGGCAGCTCCCTGACCCCGCGCTGCCCGCGCTCGGGGTAAATGTGGCAT	2593
Db	172540	CACTAGTTTTTCTAGAGAGCTGCGCCCAACTGTCTGACCCCACTAGCTACAGGGAAAGGCGAG	172481
QY	2594	TACGAGCCCAACAGAGTTTGAAGCAATCAGCTGAGACGTGGGTAAAGATGAACAGCT	2653
Db	172480	GTTGGTATCAATAGGTTAACACAGGGGCTGTGAGCTGCGCAGCTAGTTCGAACATPAACAGT	172421
QY	2654	TTAAGTTGGATTTAAGAGCTTTTAAAGATTAATCTCTGAAAGAAATGACGTA	2713
Db	172420	TTAATTTGGAATTTAAGAGGCTACACACACACACATCTGAAGAAAGATGAGTTAGTAC	172361
QY	2714	A-----CCACAGCGTACTATGAAGAGCTGTATTTATATAAAGAACGCTGGGCCATGAA	2768
Db	172360	ATTAGTACCAAGGCTGGGCTGTTTTTTGTTTCGTTTTGTATTAAGAAATCGGGCCACAAG	172301
QY	2769	CTCATTAAGCTGCAATGAG-----GTCAACATAGTATCTTTATGTAGATACTTATGATT	2820
Db	172300	CTCATATGAGACTTTTATCAATGTGACGTATGACTACATGATATGTATGTATGTACT	172241
QY	2821	ACTAAATATATATTTT 2835	
Db	172240	ACTAAATATGAAATCT 172226	
RESULT 13			
AC136661			
LOCUS			
DEFINITION	AC136661	220118 bp	DNA linear HTG 23-NOV-2002
ACCESSION	AC136661		
VERSION	AC136661.2	GI:25188309	
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_ENRICHED.		
SOURCE	Rattus norvegicus (Norway rat)		
ORGANISM	Rattus norvegicus		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;		

REFERENCE  
AUTHORS  
Schizognathi; Muroidea; Muridae; Murinae; Rattus.  
1 (bases 1 to 220118)  
Munzy,D,Marie, Metzker,M,lee, Abramson,S, Adams,C, Alder,J, Allen,C, Allen,H, Alsbrooks,S, Amin,A, Angilano,D, Anyalebech,V, Aoyagi,A, Ayodeji,M, Baca,B, Baden,H, Baldwin,D, Bandaranaike,D, Barber,M, Barnstead,M, Benahmed,F, Biewalo,K, Blair,J, Blankenburg,K, Blyth,P, Brown,M, Bryant,N, Buhay,C, Burch,P, Burrell,K, Calderon,B, Cardenas,V, Carter,K, Cavazos,I, Caesar,H, Center,A, Chacko,J, Chavez,D, Chen,G, Chen,R, Chen,Y, Chen,Z, Chu,J, Cleveland,C, Cockrell,R, Cox,C, Coyle,M, Cree,A, D'Souza,L, Davila,M,L, Davis,C, Davy-Carroll,L, De Anda,C, Dederich,D, Delgado,O, Denson,S, Deramo,C, Ding,Y, Dinh,K, Divya,K, Draper,H, Dugan-Rocha,S, Dunn,A, Durbin,K, Duval,B, Eaves,K, Egan,A, Escotto,M, Eugene,C, Evans,C,A, Falls,T, Fan,G, Fernandez,S, Finley,M, Flagg,N, Forbes,L, Foster,M, Foster,P, Frazer,C,M, Gabisi,A, Ganta,R, Garcia,A, Garner,T, Garza,M, Gebregeorgis,E, Geer,K, Gill,R, Grady,M, Guerra,W, Guevara,W, Gunaratne,P, Haaland,W, Hamil,C, Hamilton,C, Hamilton,K, Harvey,Y, Havlak,P, Hawes,A, Henderson,N, Hernandez,J, Hernandez,R, Hines,S, Hladun,S,L, Hodgson,A, Hogues,M, Hollins,B, Howells,S, Hulik,S, Hume,J, Idlebird,D, Jackson,A, Jackson,L, Jacob,L, Jiang,H, Johnson,B, Johnson,R, Jolyvet,A, Karpathy,S, Kelly,S, Kelly,S, Khan,Z, King,L, Koyar,C, Kowis,C, Kraft,C,L, Lebow,H, Levan,J, Lewis,L, Li,Z, Liu,J, Liu,J, Liu,W, Liu,Y, London,P, Longacre,S, Lopez,J, Lorenshubwa,L, Loulleged,H, Lozado,R,J, Lu,X, Ma,U, Maheshwari,M, Mahindaratne,M, Mahmoud,M, Malloy,K, Mangum,A, Mangum,B, Mapua,P, Martin,K, Martin,R, Martinez,E, Mawlinay,S, McLeod,M,P, McNeill,T,Z, Meenen,B, Milsavajevic,A, Miner,G, Minja,E, Montemayor,J, Moore,S, Morgan,M, Morris,K, Morris,S, Mundasa,M, Murphy,J, Naik,L, Nankervis,C, Neal,D, Newton,N, Nguyen,N, Norris,S, Nwackemele,O, Okunolu,G, Olamugbeogun,A, Pal,S, Parks,K, Pasernak,S, Paul,H, Perez,A, Perez,L, Pfankuch,C, Plopper,F, Polidexter,A, Popovic,D, Prims,B, Pu,L,L, Puzo,M, Quiroz,J, Rachlin,E, Reeves,K, Regier,M,A, Reich,R, Reilly,B, Reilly,M, Ren,Y, Reuter,M, Richards,S, Riggs,F, Rivers,C, Rodkey,T, Rojas,A, Rose,M, Rose,R, Ruiz,S,U, Sanders,W, Savery,G, Scherer,S, Scott,G, Shatman,S, Shan,H, Shetty,J, Shvartbeyn,A, Sison,I, Sitter,C,D, Smajls,D, Sneed,A, Sodergren,B, Song,X,Z, Sorrell,R, Soza,D, Steimle,M, Strong,R, Sutton,A, Svatek,A, Taber,P, Taylor,C, Taylor,T, Thomas,N, Thomas,S, Tingey,A, Trejos,Z, Umani,K, Vales,R, Vera,V, Villaseana,D, Waldron,L, Walker,B, Wang,J, Wang,Q, Wang,S, Warren,J, Warren,R, Wei,X, White,P, Williams,G, Willson,R, Wlarczyk,R, Wooden,H, Wolley,K, Wright,D, Wright,R, Wu,J, Yakub,S, Yen,J, Yoon,L, Yoon,V, Yu,F, Zhang,J, Zhou,J, Zhou,X, Zhao,S, Dunn,D, von Niederhuesern,A, Weiss,R, Smith,D,R, Holt,R,A, Smith,H,O, Weinstock,G, and Gibbs,R,A.  
Direct Submission  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
JOURNAL  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
JOURNAL  
TITLE  
JOURNAL  
COMMENT  
Submitted (23-NOV-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
On Nov 23, 2002 this sequence version replaced g1:24756853.  
The sequence in this assembly is a combination of BAC based reads and whole genome shotgun sequencing reads assembled using Atlas (http://www.hgsc.bcm.tmc.edu/projects/rac/). Each contig described in the feature table below represents a scaffold in the Atlas assembly (a 'contig-scaffold'). Within each contig-scaffold, individual sequence contigs are ordered and oriented, and separated by sized gaps filled with Ns to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence contigs within a contig-scaffold that consist entirely of whole genome shotgun sequence reads. Both end sequences and whole genome shotgun sequence only contigs will be indicated in the feature table.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: http://www.hgsc.bcm.tmc.edu/  
Contact: hgsc-help@bcm.tmc.edu  
----- Project Information  
Center project name: KBUP  
Center clone name: CH230-486F10  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 189341 bases at least Q40  
Consensus quality: 190385 bases at least Q30  
Consensus quality: 191057 bases at least Q20  
Estimated insert size: 195411; sum-of-contigs estimation  
Quality coverage: 6x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank\_draft\_data.html).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 1 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submittor.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
1 220118: contig of 220118 bp in length.  
Location/Qualifiers  
1..220118  
/organism="Rattus norvegicus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:10116"  
/clone="CH230-486F10"  
1..1070  
/note="wgs contig"  
211796..212627  
/note="clone boundary  
clone\_end:T7  
site:dbol  
end\_sequence:B2203379"  
215456..216816  
/note="wgs end\_extension  
clone\_end:T7"  
217092..220118  
/note="wgs end\_extension  
clone\_end:T7"  
ORIGIN  
Query Match 22.7%; Score 709.8; DB 14; Length 220118;  
Best Local Similarity 61.9%; Pred. No. 2.8e-83;  
Matches 1573; Conservative 0; Mismatches 767; Indels 203; Gaps 20;  
QY 362 GCGTGAAGCCCGCTTACCATGATTCAGAGGGGTGGACAGCTCCGCGCGGAGCAGCGGG 421  
DB 73252 GCGGGGGGGGGGGGGGGGCGAGAGGGCGGAGGAGGAGGAGCTCTCGGCTTACAGCGGGC 73311  
QY 422 TGGGGGGGTCTTACGAAAACCTTACCCGCGCGCTTGGAGGGCTTAAAGCGGAGCGGC 481  
DB 73312 TGGAGGCGGTCTTACGAAAACGATCTGTGCTTGGCGCAGAGGCTTAAAGCGCGGCGTC 73371  
QY 482 GCGTTCGACGCTTGTCCCGGAGTTGGACACCGAGAGATGGAGACCGCACCTTCA 541  
DB 73372 AGCTTGAAGACTT-----GGGTCGATACGATGGGGCTTCGACTCTGG 73415  
QY 542 GCTTCGACGAGGACCGCTGAGGCGGCGGTGCAGAGACACGACGTTGACTCGGA 601  
DB 73416 GTTTCGAGGAGCGCGCGTGAAGAAAGTTATGCAAAAGACCTTACCGTGACTCGGA 73475





Db 75518 GTTAGACTTAGTACCAAGCTGGCTTTTGTTCCTTTTGTATATAAGAAAGCTG 75577  
Qy 2761 GCCATGAACCTACCTGCGCAATGA-----GTCAACATGATATCTTATATGATATA 2812  
Db 75578 GCCACAGCTCATATGAGCTTTTACATTTAGTACGTAGTACGTACATGATATGATG 75637  
Qy 2813 CTTAGATTACTAATATATATTT 2835  
Db 75638 TGTAGCTACTAATAATATGATCT 75660

RESULT 14  
LOCUS CQ748891 708 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 34825 from Patent WO2068579.  
ACCESSION CQ748891  
VERSION CQ748891.1 GI:42376058  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1  
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
TITLE Kites, such as nucleic acid arrays, comprising a majority of  
humanecons or transcripts, for detecting expression and other uses  
thereof  
JOURNAL Patent: WO 02068579-A 34825 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
source location/Qualifiers  
1..708  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 22.6%; Score 706.4; DB 6; Length 708;  
Best Local Similarity 99.9%; Pred. No. 2.6e-82;  
Matches 707; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 858 ATGGCAGGAGGAGAGTGCAGAGCGCTGCTGACGCGGCTCAACAAGACGACTGCTGCTAC 917  
Db 1 ATGGCAGGAGGAGAGTGCAGAGCGCTGCTGACGCGGCTCAACAAGACGACTGCTGCTAC 60  
Qy 918 CACCACTGTGTCTGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 977  
Db 61 CACCACTGTGTCTGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
Qy 978 CAAAAGACGCGCAGAGAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1037  
Db 121 CAAAAGACGCGCAGAGAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
Qy 1038 GTGCTGCGGACCGGCGGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1097  
Db 181 GTGCTGCGGACCGGCGGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240  
Qy 1098 GCTTCTCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1157  
Db 241 GCTTCTCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
Qy 1158 GCGCGCTTCCGCTGACCG 1217  
Db 301 GCGCGCTTCCGCTGACCG 360  
Qy 1218 TCTTCGCGGCTGCG 1277  
Db 361 TCTTCGCGGCTGCG 420  
Qy 1278 GACTTGACGCTGCGCGGACCTGCGGAGCTGAGAGCGGAGGCTCTTCAAGGTGCGGAGATG 1337  
Db 421 GACTTGACGCTGCGCGGACCTGCGGAGCTGAGAGCGGAGGCTCTTCAAGGTGCGGAGATG 480

Qy 1338 ATCGACATGAGAGTGAAGGTCAACGTGCGCGGCTGGAACCGGCAAGCCCGGAGGCG 1397  
Db 481 ATCGACATGAGAGTGAAGGTCAACGTGCGCGGCTGGAACCGGCAAGCCCGGAGGCG 540  
Qy 1398 GCGGCGCGGAGCTCTGTCCACCGGTCAAGCGCGGCGCGCTCTGTGTGTCTTGGAG 1457  
Db 541 GCGGCGCGGAGCTCTGTCCACCGGTCAAGCGCGGCGCGCTCTGTGTGTCTTGGAG 600  
Qy 1458 GAGCGCGGCGGAGGTTGGAGACCCAGAGAGGCGCTGCGCGGCTCTTGGAGCGCG 1517  
Db 601 GAGCGCGGCGGAGGTTGGAGACCCAGAGAGGCGCTGCGCGGCTCTTGGAGCGCG 660  
Qy 1518 CTGCTGCGGCGGCTGAGCGCGCTGAGCGGCTGAGCGGAGCTGAGCTGA 1565  
Db 661 CTGCTGCGGCGGCTGAGCGCGCTGAGCGGCTGAGCGGAGCTGAGCTGA 708

RESULT 15  
LOCUS CQ746021 708 bp DNA linear PAT 03-FEB-2004  
DEFINITION Sequence 31955 from Patent WO2068579.  
ACCESSION CQ746021  
VERSION CQ746021.1 GI:42365754  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1  
AUTHORS Venter, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
TITLE Kites, such as nucleic acid arrays, comprising a majority of  
humanecons or transcripts, for detecting expression and other uses  
thereof  
JOURNAL Patent: WO 02068579-A 31955 06-SEP-2002;  
PE Corporation (NY) (US)  
FEATURES  
source location/Qualifiers  
1..708  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

ORIGIN  
Query Match 22.6%; Score 704.8; DB 6; Length 708;  
Best Local Similarity 99.7%; Pred. No. 4.2e-82;  
Matches 706; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 858 ATGGCAGGAGGAGTGCAGAGCGCTGCTGACGCGGCTCAACAAGACGACTGCTGCTAC 917  
Db 1 ATGGCAGGAGGAGTGCAGAGCGCTGCTGACGCGGCTCAACAAGACGACTGCTGCTAC 60  
Qy 918 CACCACTGTGTCTGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 977  
Db 61 CACCACTGTGTCTGACCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 120  
Qy 978 CAAAAGACGCGCAGAGAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1037  
Db 121 CAAAAGACGCGCAGAGAGCTGAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 180  
Qy 1038 GTGCTGCGGACCGGCGGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1097  
Db 181 GTGCTGCGGACCGGCGGCTGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 240  
Qy 1098 GCTTCTCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1157  
Db 241 GCTTCTCGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 300  
Qy 1158 GCGCGCTTCCGCTGACCG 1217  
Db 301 GCGCGCTTCCGCTGACCG 360  
Qy 1218 TCTTCGCGGCTGCG 1277



```
Db      ||||| 361 TCCTCCGAGGTGGCGGCGCGGCTGAGACCCGCGAGCTGCGGCTCGAGGGGAGGGC 420
QY      ||||| 1278 GACTTCGACGTGCGCGGACCTGCGGGAGCTGGAGCGGAGGTCTTCAGGTGGGCGAGATG 1337
Db      ||||| 421 GACTTCGACGTGCGGACCTGCGGGAGCTGGAGCGGAGGTCTTCAGGTGGGCGAGATG 480
QY      ||||| 1338 ATCGACACATGAGATGAAAGTCAACGTGCCCGGCTGGAACGTGCAAGCCCGCAGGCG 1397
Db      ||||| 481 ATGACACATGAGATGAAAGTCAACGTGCCCGGCTGGAACGTGCAAGCCCGCAGGCG 540
QY      ||||| 1398 GCGGGGCGGAGCTCCTGTCCAGGTCAGCGCGGCGGCCCTCTCGGTCGTCTTGCA 1457
Db      ||||| 541 GCGGGGCGGAGCTCCTGTCCAGGTCAGCGCGGCGGCCCTCTCGGTCGTCTTGCA 600
QY      ||||| 1458 GAGCGCGGGGGGGTGGCGAACCCGAGAAAGCCCTGGCGGCGCATCTTTTCGGCGCGCTG 1517
Db      ||||| 601 GAGCGCGGGGGGGTGGCGAACCCGAGAAAGCCCTGGCGGCGCATCTTTTCGGCGCGCTG 660
QY      ||||| 1518 CTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAGCTGA 1565
Db      ||||| 661 CTGCTGGCGGCTGTGGCCCTAGCCGTGTGCGTGGCGAGCTGA 708
```

Search completed: May 11, 2006, 05:54:09  
Job time : 15360 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2006, 01:07:20 ; Search time 1766 Seconds  
(without alignments)  
11782.090 Million cell updates/sec

Title: US-10-760-320A-102

Perfect score: 3122

Sequence: 1 actaagagggcgggttagcgc.....acagagcaagactctctc 3122

Scoring table:

IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4996997 seqs, 333346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N\_Geneseq\_21:\*

1: geneseqn1980s:\*  
2: geneseqn1990s:\*  
3: geneseqn2000s:\*  
4: geneseqn2001as:\*  
5: geneseqn2001bs:\*  
6: geneseqn2002as:\*  
7: geneseqn2002bs:\*  
8: geneseqn2003as:\*  
9: geneseqn2003bs:\*  
10: geneseqn2003cs:\*  
11: geneseqn2003ds:\*  
12: geneseqn2004as:\*  
13: geneseqn2004bs:\*  
14: geneseqn2005s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3122	100.0	3122	13	AD062941	AD062941 Novel hum
2	2308	73.9	2337	12	AD070462	AD070462 Pull leng
3	1009	32.3	1009	4	AAK83424	AAK83424 Human imm
4	1009	32.3	1009	4	AAK83422	AAK83422 Human imm
5	1005.8	32.2	1009	4	AAK83422	AAK83422 Human imm
6	883.2	28.3	973	4	AAK62785	AAK62785 Human imm
7	846.8	27.1	850	5	AAK93728	AAK93728 DNA encod
8	726.6	23.3	1061	6	ABQ40781	Abq40781 Oligonuc1
9	726.6	23.3	1061	6	ABQ40780	Abq40780 Oligonuc1
10	706.4	22.6	708	12	ACH87504	ACH87504 Human gen
11	594.4	19.0	1061	6	ABQ40778	Abq40778 Oligonuc1
12	594.4	19.0	1061	6	ABQ40779	Abq40779 Oligonuc1
13	524	16.8	524	12	ACH3793	ACH3793 Human gen
14	474.4	15.2	476	4	AAK83430	AAK83430 Human imm
15	474.4	15.2	476	4	AAK83427	AAK83427 Human imm
16	472.8	15.1	476	4	AAK83426	AAK83426 Human imm
17	318.6	10.2	1349	5	AAK72508	AAK72508 DNA encod
18	256.2	8.2	432	5	AAK93725	AAK93725 DNA encod
19	180.2	5.8	185371	6	ABT10718	ABT10718 Human bre

20	177.2	5.7	24295	12	ADP31706	ADP31706 Human oes
21	176.4	5.7	158001	12	ADL17884	ADL17884 Human pho
22	176	5.6	19300	12	ADP74371	ADP74371 Human X c
23	175.6	5.6	3815	12	AD085711	AD085711 Human tum
24	175	5.6	283	4	AAK86976	AAK86976 Human imm
25	174.8	5.6	22021	11	ACN44670	ACN44670 Human gen
26	174.8	5.6	73100	14	AD242280	AD242280 Human end
27	174.8	5.6	99014	6	ABN96931	ABN96931 Gene #342
28	174.8	5.6	109906	6	ABK94411	Abk94411 DNA encod
29	174.8	5.6	109906	6	ABK94411	Abk94411 DNA encod
30	174.6	5.6	65454	12	ADL08112	ADL08112 Human gen
31	174.4	5.6	110000	14	AD245062_01	AD245062_01 of
32	174.2	5.6	110000	6	ABX08336_09	Continuation (10 o
33	174.2	5.6	110000	12	ADJ25985_09	Continuation (10 o
34	174.2	5.6	110000	12	ADN97989_09	Continuation (10 o
35	174.2	5.6	110000	12	ADO50281_09	Continuation (10 o
36	174.2	5.6	110000	14	ABE85185_09	Continuation (10 o
37	174	5.6	13695	10	ADC86452	ADC86452 Human GPC
38	174	5.6	20001	14	AD259513	AD259513 Secondary
39	174	5.6	64958	14	AD213534	AD213534 Human can
40	174	5.6	65387	13	ABD33284	ABD33284 Human can
41	174	5.6	65387	13	ADR67040	ADR67040 Human can
42	174	5.6	89014	14	ABE77360	ABE77360 Human TGF
43	174	5.6	168821	11	ACN44262	ACN44262 Human gen
44	173.2	5.5	104644	6	ABQ99653	Abq99653 Human MS4
45	173.2	5.5	143306	6	ABK49586	Abk49586 Human tra

#### ALIGNMENTS

RESULT 1	
AD062941	
ID	AD062941 standard; cDNA; 3122 BP.
AC	AD062941;
XX	
DT	07-OCT-2004 (first entry)
XX	
DE	Novel human cDNA sequence #102.
XX	
XX	sa; gene; osteopathic; neuroprotective; nootropic; antiparkinsonian;
KW	cytotoxic; gene therapy; diagnostic marker; morbid state; osteoporosis;
KW	neurological disease; Alzheimer's disease; Parkinson's disease; dementia;
KW	cancer.
OS	Homo sapiens.
XX	
PN	EP1440981-A2.
XX	
PD	28-JUL-2004.
XX	
PF	21-JAN-2004; 2004EP-00001196.
XX	
PR	21-JAN-2003; 2003JP-00102206.
PR	09-MAY-2003; 2003JP-00131392.
XX	
PA	(REAS-) RES ASSOC BIOTECHNOLOGY.
XX	
PI	Isogai T, Sugiyama T, Otsuki T, Wakamatsu A, Sato H, Ishii S;
PI	Yamamoto J, Isono Y, Nagai K, Irie R;
DR	WPI, 2004-535376/52.
DR	P-PDB; AD065129.
XX	
PT	Novel 2495 cDNA, useful for treating osteoporosis, neurological diseases,
PT	Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.
XX	
PS	Claim 1; SEQ ID NO 102; 2449bp; English.
XX	
CC	The invention relates to 2495 novel polynucleotides (I) and their encoded
CC	polypeptides, sequences hybridizing to these nucleotides, sequences
CC	encoding partial polypeptides and sequences having 70% or 90% identity to

CC the nucleotide and protein sequences. The nucleotides and polypeptides  
CC are useful as diagnostic markers or therapeutic target for the diseases  
CC or morbid states. They are also useful for treating osteoporosis,  
CC neurological diseases, Alzheimer's diseases, Parkinson's diseases,  
CC dementia and various cancers. This sequence corresponds to a nucleotide  
CC sequence of the invention.

XX Sequence 3122 BP; 601 A; 891 C; 1019 G; 611 T; 0 U; 0 Other;

Query Match 100.0%; Score 3122; DB 12; Length 3122;  
Best Local Similarity 100.0%; Pred. No. 0;  
Matches 3122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACTAAGGTGGGGTTCGCTTGGAGACCCAGCACTGAGCCGCAAGCGGACGAGC 60  
DB 1 ACTAAGGTGGGGTTCGCTTGGAGACCCAGCACTGAGCCGCAAGCGGACGAGC 60  
QY 61 ACCTGACCCCGGCGGCGCCAGCCCTCGGATTGCGAGTCACTGCTCGCTTTGGGACAG 120  
DB 61 ACCTGACCCCGGCGGCGCCAGCCCTCGGATTGCGAGTCACTGCTCGCTTTGGGACAG 120  
QY 121 GAGGTGCCAGTCTCTGCGGGGCAACCCGAGTCTCTGTGCGCAAGAGGTCCGGAGTCACT 180  
DB 121 GAGGTGCCAGTCTCTGCGGGGCAACCCGAGTCTCTGTGCGCAAGAGGTCCGGAGTCACT 180  
QY 181 ATAGCTGGGTTCTAATGCCATCAACAGCAAAAATCTCCGCGGAGCTTGGCCGCTTTTAA 240  
DB 181 ATAGCTGGGTTCTAATGCCATCAACAGCAAAAATCTCCGCGGAGCTTGGCCGCTTTTAA 240  
QY 241 CCTGGGCTCAATTTCCCACTCGTAAATAGAAAGGGATTGAGTCTCCGAGGCTTAA 300  
DB 241 CCTGGGCTCAATTTCCCACTCGTAAATAGAAAGGGATTGAGTCTCCGAGGCTTAA 300  
QY 301 TTCGAAGACTCGGATGAGGCGAAAGGAGAGAGATGAGGCCACCAAGTGAAGCTCTCC 360  
DB 301 TTCGAAGACTCGGATGAGGCGAAAGGAGAGAGATGAGGCCACCAAGTGAAGCTCTCC 360  
QY 361 CCGGTGAGACCCCGGCTTACCTGATCCAGGGGTGAGCTCCGCGCGGAGAGAGCGG 420  
DB 361 CCGGTGAGACCCCGGCTTACCTGATCCAGGGGTGAGCTCCGCGCGGAGAGAGCGG 420  
QY 421 GTGAGCGGGTCTTAAGAAACCTTACCCGCGGCTTGGCAGCGGCTTAAAGCGAGCGG 480  
DB 421 GTGAGCGGGTCTTAAGAAACCTTACCCGCGGCTTGGCAGCGGCTTAAAGCGAGCGG 480  
QY 481 CGGCTCTGACGCTTGTCCCGGAGTTGGCAACGAGAGATGGGACCGCACCTTC 540  
DB 481 CGGCTCTGACGCTTGTCCCGGAGTTGGCAACGAGAGATGGGACCGCACCTTC 540  
QY 541 AGCTTGGCAGGAGCCACCGTGAAGGCGGCGGTGGCAGAGCAAGACGTGTGATTCGG 600  
DB 541 AGCTTGGCAGGAGCCACCGTGAAGGCGGCGGTGGCAGAGCAAGACGTGTGATTCGG 600  
QY 601 AGTGCCTTGGGAGAGTGAACGAGGAGCGGGGAGCCGCTAAAGGGCTCCCTCTGCGC 660  
DB 601 AGTGCCTTGGGAGAGTGAACGAGGAGCGGGGAGCCGCTAAAGGGCTCCCTCTGCGC 660  
QY 661 GCGCCGTCGAGAGGCGCAGTGAAGGTCCCGGCGGAGCTCCGTGAGAGTTGGCGGTA 720  
DB 661 GCGCCGTCGAGAGGCGCAGTGAAGGTCCCGGCGGAGCTCCGTGAGAGTTGGCGGTA 720  
QY 721 GCGCCGAGAGAGTCAAGGACATGAAGAGGTTCTGTCGCGCGGAGCCAAAGCCGAGAG 780  
DB 721 GCGCCGAGAGAGTCAAGGACATGAAGAGGTTCTGTCGCGCGGAGCCAAAGCCGAGAG 780  
QY 781 GGGGTTAGCCATCTCTGCGGCTGAGGGGAGGCTTAAAGGCGCGGCGCGCGGCC 840  
DB 781 GGGGTTAGCCATCTCTGCGGCTGAGGGGAGGCTTAAAGGCGCGGCGCGCGGCC 840  
QY 841 AGCGGAGCCACCGGATGAGGAGGAGAGAGTGAAGGCGCTGTGAGACGGCTTACA 900  
DB 841 AGCGGAGCCACCGGATGAGGAGGAGAGAGTGAAGGCGCTGTGAGACGGCTTACA 900  
QY 901 AGCGGAGCCACCGGATGAGGAGGAGAGAGTGAAGGCGCTGTGAGACGGCTTACA 960  
DB 901 AGCGGAGCCACCGGATGAGGAGGAGAGAGTGAAGGCGCTGTGAGACGGCTTACA 960  
QY 961 ACTTCGCGAGAGCTGCAAAAGAGCGGCGAGAGAGAGAGTGGCGGTGCACCT 1020  
DB 961 ACTTCGCGAGAGCTGCAAAAGAGCGGCGAGAGAGAGAGTGGCGGTGCACCT 1020  
QY 1021 GCGCCGCTGACCTGCTGTGTCGCGGACCGGAGGCTTGGCGGCGGAGCCGAGT 1080  
DB 1021 GCGCCGCTGACCTGCTGTGTCGCGGACCGGAGGCTTGGCGGCGGAGCCGAGT 1080  
QY 1081 TCGAGGAGCTGAGGAGGCTTCTGCGGAGCTGAGACCTGCTGAGAGGAGCATGCGAC 1140  
DB 1081 TCGAGGAGCTGAGGAGGCTTCTGCGGAGCTGAGACCTGCTGAGAGGAGCATGCGAC 1140  
QY 1141 GCTCGTAGAGCTGAGGAGCGCGGCTTCCGCTTGAACGCGCGGAGCCGCTGTGCGCA 1200  
DB 1141 GCTCGTAGAGCTGAGGAGCGCGGCTTCCGCTTGAACGCGCGGAGCCGCTGTGCGCA 1200  
QY 1201 CAGGTGTGCTGCGGCTTCTCGGCGTGGCGGCGGCGGCTTGAACCCGACCTTGC 1260  
DB 1201 CAGGTGTGCTGCGGCTTCTCGGCGTGGCGGCGGCGGCTTGAACCCGACCTTGC 1260  
QY 1261 GCGTGGGCGGAGGAGGAGCTTGAAGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1320  
DB 1261 GCGTGGGCGGAGGAGGAGGAGCTTGAAGTGGAGGAGGAGGAGGAGGAGGAGGAGG 1320  
QY 1321 TTCAAGTGGGCGAGTGAATGCAAAATGAGATGAGAGTGAAGTGAAGTGAAGTGAAG 1380  
DB 1321 TTCAAGTGGGCGAGTGAATGCAAAATGAGATGAGAGTGAAGTGAAGTGAAGTGAAG 1380  
QY 1381 TGAAGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440  
DB 1381 TGAAGCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1440  
QY 1441 CGGTGTGCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500  
DB 1441 CGGTGTGCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1500  
QY 1501 TCGTTTTCGCGCGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560  
DB 1501 TCGTTTTCGCGCGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1560  
QY 1561 GCTGACAGACCCGAGCGGCGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620  
DB 1561 GCTGACAGACCCGAGCGGCGGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1620  
QY 1621 GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680  
DB 1621 GATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1680  
QY 1681 GTACACGCGGCTTCAAGTGAACATCTGCTGGGAGAGACAGGTTTCTCTTCTGCGC 1740  
DB 1681 GTACACGCGGCTTCAAGTGAACATCTGCTGGGAGAGACAGGTTTCTCTTCTGCGC 1740  
QY 1741 CCGGAGAGAGTAACTTGGCGCGGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800  
DB 1741 CCGGAGAGAGTAACTTGGCGCGGCGGCTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1800  
QY 1801 TTAATTCCTAATTAAGAAACCGTCAAGTGAACCTAGATCCCTCCGAGTTAAGAGTT 1860  
DB 1801 TTAATTCCTAATTAAGAAACCGTCAAGTGAACCTAGATCCCTCCGAGTTAAGAGTT 1860  
QY 1861 AACATGTGCTGTGGGCGCTTTTACAGGAGTCCGAGTTCCGAGCCACCCCTGCGCA 1920  
DB 1861 AACATGTGCTGTGGGCGCTTTTACAGGAGTCCGAGTTCCGAGCCACCCCTGCGCA 1920  
QY 1921 GCGTGGCCCTTTCTGCGTGGAGAGATTGAAGAGGAGGAGGAGGAGGAGGAGGAG 1980  
DB 1921 GCGTGGCCCTTTCTGCGTGGAGAGATTGAAGAGGAGGAGGAGGAGGAGGAGGAG 1980  
QY 1981 GAGAGGAGCGTGTGTTCTATGTGTTGCTGTTCCGAGCAAGAAAAATTGCA 2040

Db	19b1	GAGAGGAGCGCTGTTGGTTCTATGTGGTGGTCTGTTTCCCGACAAAGAAAATTGCA	2040
Qy	2041	TCAATGTCAAGAGCTTTATTACTTAATCTTTACAGGGCCCTAAATTAGAGATGTCC	2100
Db	2041	TCAATGTCAAGAGCTTTATTACTTAATCTTTACAGGGCCCTAAATTAGAGATGTCC	2100
Qy	2101	TGAAGGAGTTCATACAAAGGGCTTCTCTAAGAGCCCTACAGCCCTCTTAGAGAGT	2160
Db	2101	TGAAGGAGTTCATACAAAGGGCTTCTCTAAGAGCCCTACAGCCCTCTTAGAGAGT	2160
Qy	2161	TTATTCATTCGTCCTCCCAAGAGAGCTTAAGAGATTTAGAGTCATGACCTCCACATGCG	2220
Db	2161	TTATTCATTCGTCCTCCCAAGAGAGCTTAAGAGATTTAGAGTCATGACCTCCACATGCG	2220
Qy	2221	CTCAGGGGCTGACCTTATTTAGAAACCAAGAGGGGTGAACCTTACTCACGAGC	2280
Db	2221	CTCAGGGGCTGACCTTATTTAGAAACCAAGAGGGGTGAACCTTACTCACGAGC	2280
Qy	2281	TTTGGATCCAGTGCAGCACCTTGCTCTCGGAAAAGGGCTCTCCCAAGCACCCGAGATGG	2340
Db	2281	TTTGGATCCAGTGCAGCACCTTGCTCTCGGAAAAGGGCTCTCCCAAGCACCCGAGATGG	2340
Qy	2341	GGGTAAAGAGAAAGCAGAGGCTTGGGGTAAAGGCCACCTGGATTTAAACAGGACCTTC	2400
Db	2341	GGGTAAAGAGAAAGCAGAGGCTTGGGGTAAAGGCCACCTGGATTTAAACAGGACCTTC	2400
Qy	2401	TCTTCTCTGGGGCTTATTTTGTTCAGAACTAGACAGAGTGTGTGAACCTCTTGTGCA	2460
Db	2401	TCTTCTCTGGGGCTTATTTTGTTCAGAACTAGACAGAGTGTGTGAACCTCTTGTGCA	2460
Qy	2461	GAAGGGCTGGGAATCTCTTTAAGACACTTAATCCTATTTATCCCTGGAATGTGCTGC	2520
Db	2461	GAAGGGCTGGGAATCTCTTTAAGACACTTAATCCTATTTATCCCTGGAATGTGCTGC	2520
Qy	2521	TGGCAGATGAGAGGGCTGGCTTTGGCAGCTCCGACCCCGCGTGGCCGCCCTCCG	2580
Db	2521	TGGCAGATGAGAGGGCTGGCTTTGGCAGCTCCGACCCCGCGTGGCCGCCCTCCG	2580
Qy	2581	GGTATGTGGCATTACTGCGCCACAGAGTTTYTGACCAATCAGCTCGAGACTGGGTTA	2640
Db	2581	GGTATGTGGCATTACTGCGCCACAGAGTTTYTGACCAATCAGCTCGAGACTGGGTTA	2640
Qy	2641	GAAATGTAAACAGCTTTAATTGGGATTTAAGAACTTTTAAAGTATATCTCTGAA	2700
Db	2641	GAAATGTAAACAGCTTTAATTGGGATTTAAGAACTTTTAAAGTATATCTCTGAA	2700
Qy	2701	GAAAAATGACGTAAACCAAGCGTGAACATAGAGCTGTATTTTAAATAAGAAAGCGTGG	2760
Db	2701	GAAAAATGACGTAAACCAAGCGTGAACATAGAGCTGTATTTTAAATAAGAAAGCGTGG	2760
Qy	2761	GCCATGAACTCATACCTGCAATGAGTCAAACTATGATCTTTATATGATACCTTAAGTT	2820
Db	2761	GCCATGAACTCATACCTGCAATGAGTCAAACTATGATCTTTATATGATACCTTAAGTT	2820
Qy	2821	ACTAATATATATTTCACTACTTCTGAGGTGATATGTCCTTCCGCCCCGCCCACTTTT	2880
Db	2821	ACTAATATATATTTCACTACTTCTGAGGTGATATGTCCTTCCGCCCCGCCCACTTTT	2880
Qy	2881	TCTTTTTGAGGCAAGTGGATCACTTGAGGCCAGAGATTTGAGACAGACCTGGCCAACT	2940
Db	2881	TCTTTTTGAGGCAAGTGGATCACTTGAGGCCAGAGATTTGAGACAGACCTGGCCAACT	2940
Qy	2941	AGCGAAACCCGATCTCTACTAAAAAATTAATAATTTGCGCGGACATGCTGGCGATGCT	3000
Db	2941	AGCGAAACCCGATCTCTACTAAAAAATTAATAATTTGCGCGGACATGCTGGCGATGCT	3000
Qy	3001	GTTGTCTCCAGCTACTCGGAGGTTGAGCAGAGAGTGTGCTTGAATGCAAGAGGTGAGG	3060
Db	3001	GTTGTCTCCAGCTACTCGGAGGTTGAGCAGAGAGTGTGCTTGAATGCAAGAGGTGAGG	3060
Qy	3061	TTGCAATTAGCAAGATTGTGCACTGCACTCAGCTGGGCAACAGACAAATCTGTCT	3120
Db	3061	TTGCAATTAGCAAGATTGTGCACTGCACTCAGCTGGGCAACAGACAAATCTGTCT	3120

Db	3061	TTGCAATGAGCAAGATTGTGCACTGCACCTCCGAGGCAACAGAGCAAGACTGTTC	3122
Qy	3121	TC	3122
Db	3121	TC	3122
RESULT 2			
ID	ADRO7462	standard; cDNA; 2327 BP.	
AC	ADRO7462;		
XX			
DT	04-NOV-2004	(first entry)	
XX			
De	Full length human cDNA useful for treating neurological disease Seg 968.		
XX			
KW	gene; ss; human; oligo-capping method; diagnostic marker; gene therapy;		
KW	osteoporosis; neurological disease; Alzheimer's disease;		
KW	Parkinson's disease; dementia; short memory; cancer;		
KW	sense or motor function; emotional reaction; fear response; panic;		
KW	osteopathic; neuroprotective; nootropic; antiparkinsonian; cyostatic;		
XX	tranquilliser.		
XX			
OS	Homo sapiens.		
XX			
PN	EP1447413-A2.		
XX			
PD	18-AUG-2004.		
XX			
XP	12-FEB-2004; 2004EP-00003145.		
XX			
XP	14-FEB-2003; 2003JP-00102207.		
PR	09-MAY-2003; 2003JP-00131452.		
XX			
PA	(REAS-) RES ASSOC BIOTECHNOLOGY.		
XX			
PI	Iocgal T, Yamamoto J, Nishikawa T, Isono Y, Sugiyama T, Otsuki T,		
PI	Wakamatsu A, Ishii S, Nagai K, Irie R;		
XX			
DR	WPI; 2004-583265/57.		
XX			
DR	P-PSDB; ADRO9418.		
XX			
XX	New 1995 cDNA, useful for treating osteoporosis, neurological diseases,		
FT	Alzheimer's diseases, Parkinson's diseases, dementia and various cancers.		
XX			
PS	Claim 1; SEQ ID NO 968; 2686pp; English.		
XX			
CC	This invention relates to novel, isolated full length human cDNA		
CC	molecules and the encoded proteins thereof. Specifically, it refers to		
CC	cDNA clones obtained by an oligo-capping method, where none of these		
CC	clones are identical to any known human mRNAs. The present invention		
CC	describes an immunoassay to identify agonists and antagonists, as well as		
CC	antibodies, antisense molecules and siRNAs that can all be used to bind		
CC	to and modulate expression of the cDNA molecules. As such, these		
CC	molecules are useful for diagnostic markers or therapeutic targets for		
CC	the various diseases or morbid states. In particular, they are useful in		
CC	gene therapy for treating osteoporosis, neurological disease, Alzheimer's		
CC	disease, Parkinson's disease, dementia, short memory and various cancers,		
CC	as well as for maintaining equilibrium of sense or motor function, and		
CC	for treating emotional reaction, fear response and panic. Accordingly,		
CC	they exhibit osteopathic, neuroprotective, nootropic, antiparkinsonian,		
CC	cyostatic and tranquilliser activities. This polynucleotide is a full		
CC	length human cDNA sequence of the invention. NOTE: This sequence is not		
CC	given in the sequence listing of the specification but can be obtained on		
CC	CD-ROM from the European Patent Office, Vienna Sub-office.		
XX			
Qx	Sequence 2327 BP; 424 A; 667 C; 788 G; 448 T; 0 U; 0 Other;		
Query Match 73.9%; Score 2308; DB 13; Length 2327;			
Best Local Similarity 99.7%; Pred. No. 0;			
Matches 2322; Conservative 0; Mismatches 5; Indels 1; Gaps 1;			

QY 444 ACCGCGCGCGCTTGGCAGCGCTTAAAGCGAGCGCGCGCTCTGACGCTTGCCCC 503  
Db 1 ACCCGCGCGCGCTTGGCAGCGCTTAAAGCGAGCGCGCGCTCTGAGCTTCTTGCCCC 60  
QY 504 GGAAGTTGGCACCACGAGAGATGGGAGCCGCAACCTCAGCTTTCGAGGAGGCGCACGTTGG 563  
Db 61 GGAAGTTGGCACCACGAGAGATGGGAGCCGCAACCTCAGCTTTCGAGGAGGCGCACGTTGG 120  
QY 564 AGGCGAGGCGCGGTGCAAGACACAGCTGTACTGTGAAGTGGCGCTTGGGAGAGATGAGCG 623  
Db 121 AGGCGAGGCGCGGTGCAAGACACAGCTGTACTGTGAAGTGGCGCTTGGGAGAGATGAGCG 180  
QY 624 AGGAGCGCGGAGACCGCTTAAACGCGGCTCCCTCTGCGCGCGCGCTCGCAGAGCGGACGT 683  
Db 181 AGGAGCGCGGAGACCGCTTAAACGCGGCTCCCTCTGCGCGCGCGCTCGCAGAGCGGACGT 240  
QY 684 CGAGGAGTCCCGGAGCGGCTCCGTTGACGTTGGCGGTAGCGCGAGCGAGTCAAGACAT 743  
Db 241 CGAGGAGTCCCGGAGCGGCTCCGTTGACGTTGGCGGTAGCGCGAGCGAGTCAAGACAT 300  
QY 744 GAAAGACGTTCTGTCCGCGCGCGCCAAAGGCGGAGTGGGGGTTTACCACTCTGCGCGCG 803  
Db 301 GAAAGACGTTCTGTCCGCGCGCGCCAAAGGCGGAGTGGGGGTTTACCACTCTGCGCGCG 360  
QY 804 CTGAGGCGGAGGCTTAAACGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 863  
Db 361 CTGAGGCGGAGGCTTAAACGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420  
QY 864 AGGAGAGAGTGCAGAGCGCGCTGTGAGACGGGCTTCAACAAAGACGATCGGTCTTACCAACAC 923  
Db 421 AGGAGAGAGTGCAGAGCGCGCTGTGAGACGGGCTTCAACAAAGACGATCGGTCTTACCAACAC 480  
QY 924 CTGGGTGCTGACCGTGGGTGGCTCTCGCGCGGACCTTGGCAGAACCTTGGCGGAGAGAGCTTCAAAAG 983  
Db 481 CTGGGTGCTGACCGTGGGTGGCTCTCGCGCGGACCTTGGCAGAACCTTGGCGGAGAGAGCTTCAAAAG 540  
QY 984 AGCGCCAGAAAGCGCAGGAGCTGGCGGTGTCCACTTGGCGCGCGGTGAGCGCTTGGGTGGCGCTTTC 1043  
Db 541 AGCGCCAGAAAGCGCAGGAGCTGGCGGTGTGTCTCACTTGGCGCGCGGTGAGCGCTTGGGTGGCGCTTTC 600  
QY 1044 CGCGACCGGGGCGCTTGGCGCGCGCGACGAGCGCGCGAGTTTGAAGCGGCTTGGGTGGCGCTTTC 1103  
Db 601 CGCGACCGGGGCGCTTGGCGCGCGCGACGAGCGCGCGAGTTTGAAGCGGCTTGGGTGGCGCTTTC 660  
QY 1104 TGGGGCTGTGCTGAGACTTGTGAGAGCGGACATGCGACGCTCGTGAAGTGGGGCGCGCG 1163  
Db 661 TGGGGCTGTGCTGAGACTTGTGAGAGCGGACATGCGACGCTCGTGAAGTGGGGCGCGCG 720  
QY 1164 TTCCGCGTGCAGCGCGCGCGCGCGCGCGCTGTGCGGACAGGTGTGGCTTGGCGCTTCTTC 1223  
Db 721 TTCCGCGTGCAGCGCGCGCGCGCGCGCGCTGTGCGGACAGGTGTGGCTTGGCGCTTCTTC 780  
QY 1224 GCGGTGGCGCGCGCGCGCTGAGCACCCGCGACCTTGGCGCTTGAAGCGGAGCGGCGGACTTTC 1283  
Db 781 GCGGTGGCGCGCGCGCGCTGAGCACCCGCGACCTTGGCGCTTGAAGCGGAGCGGCGGACTTTC 840  
QY 1284 GACGTGCGCGACCTTGGGAGAGCTGAGAGCGCGAGTCTTCAAGTGGGCGGAGATGATCGAC 1343  
Db 841 GACGTGCGCGACCTTGGGAGAGCTGAGAGCGCGAGTCTTCAAGTGGGCGGAGATGATCGAC 900  
QY 1344 AACCATGAGATGAAGTCAACGTTGCCGCTGGAACGTTGAAGCGCGGAGCGCGCGCGCG 1403  
Db 901 AACCATGAGATGAAGTCAACGTTGCCGCTGGAACGTTGAAGCGCGGAGCGCGCGCGCGCG 960  
QY 1404 GCGGAGCTCTGTCAACGCTCAGCGCGCGCGCGCGCGCTCTCTGCGGTGTCTTTCAGAGAGCGCG 1463  
Db 961 GCGGAGCTCTGTCAACGCTCAGCGCGCGCGCGCGCGCTCTCTGCGGTGTCTTTCAGAGAGCGCG 1020  
QY 1464 GGGGGGGGGTTTGGACCCCGAGAGAGGCGCTGGCGCGACATCTTTTGGCGCGCGGTGGTGG 1523  
Db 1021 GGGGGGGGGTTTGGACCCCGAGAGAGGCGCTGGCGCGACATCTTTTGGCGCGCGGTGGTGG 1080  
QY 1524 GCGGCTGTGGCCTTACGCTGTGCGGTGGAGGCTGACAGACACCGGACGCGCGCG 1583

Db 1081 GCGGCTGTGGCCTTACGCTGTGCGGTGGAGAGCTGAGCTGACGAGACACCGGACGCGCGCG 1140  
QY 1584 CTGCTGTGCGCGCTTCCCTCCCTGAGAAAGACTTGGAGTGGGTGGGTCTTGGCGCTTGT 1643  
Db 1141 CTGCTGTGCGCGCTTCCCTCCCTGAGAAAGACTTGGAGTGGGTGGGTCTTGGCGCTTGT 1200  
QY 1644 GCAAGGGAGAGTGTCTTAAACCCCGGTGTGCAATGGGTATACCGGCGCTTCCATGTGAC 1703  
Db 1201 GCAAGGGAGAGTGTCTTAAACCCCGGTGTGCAATGGGTATACCGGCGCTTCCATGTGAC 1260  
QY 1704 ATCTGCTGTGGGAGGACACGCTTTCCTCTTGTGCGCGCGGAGAAAGTTAACTTTCGCGC 1763  
Db 1261 ATCTGCTGTGGGAGGACACGCTTTCCTCTTGTGCGCGCGGAGAGATTACTTTCGCGC 1319  
QY 1764 GCGCGTCAAGGACATTAACCGCTTAAACGCTGTGAGAGACTTATTCCTTATTAATGAAAAAC 1823  
Db 1320 GCGCGTCAAGGACATTAACCGCTTAAACGCTGTGAGAGACTTATTCCTTATTAATGAAAAAC 1379  
QY 1824 GTTCAAGTGAACCTTGAATCCCTCCGAGTAAATGATTAACACATGCGCTGTGGGCGGT 1883  
Db 1380 GTTCAAGTGAACCTTGAATCCCTCCGAGTAAATGATTAACACATGCGCTGTGGGCGGT 1439  
QY 1884 TTTTACAGGAGTCCGAGTTCCGATGCCACCCCTGCGACGCTGCGCCCTTCTGCGTGG 1943  
Db 1440 TTTTACAGGAGTCCGAGTTCCGATGCCACCCCTGCGACGCTGCGCCCTTCTGCGTGG 1499  
QY 1944 ACAATTGAAAAAGTGGGTGGGTGAGTGAAGTTTGAAGAGGACGCTGTTTGGTTCTA 2003  
Db 1500 ACAATTGAAAAAGTGGGTGGGTGAGTGAAGTTTGAAGAGGACGCTGTTTGGTTCTA 1559  
QY 2004 TGTGGTGTGCTGTTTCCCGGACAGAAAAATTTGCAATCAATGTACAGAGCTTATTA 2063  
Db 1560 TGTGGTGTGCTGTTTCCCGGACAGAAAAATTTGCAATCAATGTACAGAGCTTATTA 1619  
QY 2064 CTTTATCTTTCAGGCGCTTAAATTTAGAGAGTGTCTGAGAGCACTTCAACAAAGGCG 2123  
Db 1620 CTTTATCTTTCAGGCGCTTAAATTTAGAGAGTGTGTCCGAGAGGCACTTCAACAAAGGCG 1679  
QY 2124 TTTTCTTAAAGCGGCTTAAAGCGCTTCTTGAAGAGTTATTCATTCCTCCCGAAGAC 2183  
Db 1680 TTTTCTTAAAGCGGCTTAAAGCGCTTCTTGAAGAGTTATTCATTCCTCCCGAAGAC 1739  
QY 2184 GTTGAAGAGATTGAGGTCAATGACCTCCACCTGCGCGTCAAGGGGCTGACCTTATTTAG 2243  
Db 1740 GTTGAAGAGATTGAGGTCAATGACCTCCACCTGCGCGTCAAGGGGCTGACCTTATTTAG 1799  
QY 2244 AAACCAAGAGGCTGTGAACTTCTCAACGACTTGAATCCAGTGTGCGGACACTTTC 2303  
Db 1800 AAACCAAGAGGCTGTGAACTTCTCAACGACTTGAATCCAGTGTGCGGACACTTTC 1859  
QY 2304 CTGCGGAAAAAGGCTCTCCCGAGGACCCCGAGATGGGGGTAAAGAGAAAGACAGGCT 2363  
Db 1860 CTGCGGAAAAAGGCTCTCCCGAGGACCCCGAGATGGGGGTAAAGAGAAAGACAGGCT 1919  
QY 2364 TGGGGTGAAGGACCTTGGGTAAACAGGCACTTTCCTCTCTGAGGAGCTTATTTTGG 2423  
Db 1920 TGGGGTGAAGGACCTTGGGTAAACAGGCACTTTCCTCTCTGAGGAGCTTATTTTGG 1979  
QY 2424 TTCAAGAACTGACACAGAGTGTGAACTCTTTCAGAGAGGCTTGGAAATCTCTTTAG 2483  
Db 1980 TTCAAGAACTGACACAGAGTGTGAACTCTTTCAGAGAGGCTTGGAAATCTCTTTAG 2039  
QY 2484 AGCACTTAATCTTATTTATCCCTGGAATGTGCGGTGTGCGGAGTGAAGAGGCTTGGCTT 2543  
Db 2040 AGCACTTAATCTTATTTATCCCTGGAATGTGCGGTGTGCGGAGTGAAGAGGCTTGGCTT 2099  
QY 2544 GGCAGCTCCCTGAACCCCGCGGTGCGCGCGCTCCGAGGATTAATGAGCACTTACTGAGCCA 2603  
Db 2100 GGCAGCTCCCTGAACCCCGCGGTGCGCGCGCTCCGAGGATTAATGAGCACTTACTGAGCCA 2159  
QY 2604 CAGAGTTTGAAGCAATCAGCTTGAAGCTGGGTTTGAATGAATGAACAGCTTAACTTGG 2663



Db 2160 CAGAGGTTTGAACCATGCTCTGAGACTGGGTTAGATGACGTTTAATTGGG 2219  
Qy 2664 ATTTAGAGCTTTTAAAGGTATATCTCTGAAAGAAATGACGTAAACAGACGCT 2723  
Db 2220 ATTTAAGAAAGCTTTAAAGGTATATCTCTGAAAGAAATGACGTAAACAGACGCT 2279  
Qy 2724 GTACTATGAAGCTGTTATTTTATATAAGAACCTGGGCGCATGACTC 2771  
Db 2280 GTACTATGAAGCTGTTATTTTATATAAGAACCTGGGCGCATGACTC 2327

RESULT 3  
AAK83424/c  
ID AAK83424 standard; DNA; 1009 BP.  
XX AAK83424;  
AC  
DT 07-NOV-2001 (first entry)  
-XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38236.  
XX  
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
XX cytostatic; gene therapy; vaccine; metastasis; ds.  
OS Homo sapiens.  
XX  
XX W0200157182-A2.  
PN  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001W0-US001354.  
XX  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226688P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234997P.  
PR 25-SEP-2000; 2000US-0234998P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235834P.  
PR 27-SEP-2000; 2000US-0235835P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239335P.  
PR 13-OCT-2000; 2000US-0239337P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241221P.  
PR 20-OCT-2000; 2000US-0241785P.  
PR 20-OCT-2000; 2000US-0241786P.  
PR 20-OCT-2000; 2000US-0241787P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 01-NOV-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.  
PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.





PT useful for preventing, diagnosing and/or treating cancers and metastasis.  
XX Discloure: SEQ ID NO 38235; 3071bp + Sequence Listing; English.  
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAM82170 to AAM91921. (I) have cytotoxic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK81694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention  
XX  
XX Sequence 1009 BP; 260 A; 269 C; 239 G; 241 T; 0 U; 0 Other;

Query Match 32.3%; Score 1009; DB 4; Length 1009;  
Best Local Similarity 100.0%; Pred. No. 5,7e-181;  
Matches 1009; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1766 CCGTCAGGGGCAATTACCGCTACGAGAGCTTATCCCTAATTAAGAAAACCGT 1825  
DB 1009 CCGTCAGGGGCAATTACCGCTACGAGAGCTTATCCCTAATTAAGAAAACCGT 950  
QY 1826 CACAGTACCGCTAGATCCCTCCGAGTTATGATTAACAACATGCTGTGGGGGCTTT 1885  
DB 949 CACAGTACCGCTAGATCCCTCCGAGTTATGATTAACAACATGCTGTGGGGGCTTT 890  
QY 1886 TACAGGAGATCCGAGTCCGTCGCCACCCCTGCCAGCTCGCCCTTCTCGTGGAGC 1945  
DB 889 TACAGGAGATCCGAGTCCGTCGCCACCCCTGCCAGCTCGCCCTTCTCGTGGAGC 830  
QY 1946 AGTTGAAAAGGTGGGTGGGTGAGTGAAGTTTGGAGAGGAGCGCTGTTTGTCTATG 2005  
DB 829 AGTTGAAAAGGTGGGTGGGTGAGTGAAGTTTGGAGAGGAGCGCTGTTTGTCTATG 770  
QY 2006 TGGTGTGCTGTTTCCCGGCAAGAAAATTTGCAATATGTCAGACCTTTATTAAC 2065  
DB 769 TGGTGTGCTGTTTCCCGGCAAGAAAATTTGCAATATGTCAGACCTTTATTAAC 710  
QY 2066 TTAATCTTCAAGGCTTAATATTAGAGAGTGTCTGAGAGCAGTTCAATCAAGGGCTT 2125  
DB 709 TTAATCTTCAAGGCTTAATATTAGAGAGTGTCTGAGAGCAGTTCAATCAAGGGCTT 650  
QY 2126 TCTTAAAGCGCGCTACAGCCCTTCTAGCAGAGTTATCAATTCGTCCTCAAGAGCAGC 2185  
DB 649 TCTTAAAGCGCGCTACAGCCCTTCTAGCAGAGTTATCAATTCGTCCTCAAGAGCAGC 590  
QY 2186 TGAAGAGATTGAAGTCAATGACCTCCCACTGCGCTGAGGGGCTGACCTTTATAGAA 2245  
DB 589 TGAAGAGATTGAAGTCAATGACCTCCCACTGCGCTGAGGGGCTGACCTTTATAGAA 530  
QY 2246 ACCAAAGAGGGTGGTGAACCTACTCTACAGCACTTGATCAAGTGCACACTTGCT 2305  
DB 529 ACCAAAGAGGGTGGTGAACCTACTCTACAGCACTTGATCAAGTGCACACTTGCT 470  
QY 2306 GGGGAAAAGGGCTCTCCCAAGCACCAGAGATGGGGTTAAGAGAAAGCAGAGGCTTG 2365  
DB 469 GGGGAAAAGGGCTCTCCCAAGCACCAGAGATGGGGTTAAGAGAAAGCAGAGGCTTG 410  
QY 2366 GGGTAGGGCACTGCTGTTTAAACAGGCACTTCTCTCTGAGGGCTTATTTTGT 2425  
DB 409 GGGTAGGGCACTGCTGTTTAAACAGGCACTTCTCTCTGAGGGCTTATTTTGT 350  
QY 2426 CAGAACTAGACAGAGTGTGAACTCTTCTGACAGAGGGCTGGAAATCTCTTATAG 2485

DB 349 CAGAACTAGACAGAGTGTGAACTCTTCTGACAGAGGGCTGGAAATCTCTTATAG 290  
QY 2486 CACTTAATCTTATTTATCCCTGGAATGCGTGTGCGCCAGTAGAGGGCTGGCTTTGG 2545  
DB 289 CACTTAATCTTATTTATCCCTGGAATGCGTGTGCGCCAGTAGAGGGCTGGCTTTGG 230  
QY 2546 CAGCTCCCTGACCCCGGGCTGCGCCCTCCGGGGTAATGTGGCATTACTGGCCACA 2605  
DB 229 CAGCTCCCTGACCCCGGGCTGCGCCCTCCGGGGTAATGTGGCATTACTGGCCACA 170  
QY 2606 GAGGTTTGAACCAATCAGCTCTGAGACTGAGTGAATGTAAACAGCTTTAACTTGGAT 2665  
DB 169 GAGGTTTGAACCAATCAGCTCTGAGACTGAGTGAATGTAAACAGCTTTAACTTGGAT 110  
QY 2666 TTAAGAGCTTTTAAAGGTAAATTAATCTCTGAAAAGAAAATGACGTAAACAGCGCT 2725  
DB 109 TTAAGAGCTTTTAAAGGTAAATTAATCTCTGAAAAGAAAATGACGTAAACAGCGCT 50  
QY 2726 ACTATGAAGCTGTATTTTAAATTAAGAAAGCTGGGCGCATGAATCATTA 2774  
DB 49 ACTATGAAGCTGTATTTTAAATTAAGAAAGCTGGGCGCATGAATCATTA 1

RESULT 5  
AAK83422/C  
ID AAK83422 standard; DNA; 1009 BP.  
XX AAK83422;  
AC 07-NOV-2001 (first entry)  
DT  
XX  
DE Human immune/haematopoietic antigen genomic sequence SEQ ID NO:38234.  
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ds.  
XX Homo sapiens.  
OS  
PN WO200157182-A2.  
XX  
PD 09-AUG-2001.  
XX  
PF 17-JAN-2001; 2001WO-US001354.  
PR 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184644P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190076P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-020515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216807P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 11-JUL-2000; 2000US-0217496P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225447P.  
PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226681P.  
PR 22-AUG-2000; 2000US-0226686P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0229287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0232397P.  
PR 14-SEP-2000; 2000US-0232398P.  
PR 14-SEP-2000; 2000US-0232399P.  
PR 14-SEP-2000; 2000US-0232400P.  
PR 14-SEP-2000; 2000US-0232401P.  
PR 14-SEP-2000; 2000US-0233063P.  
PR 14-SEP-2000; 2000US-0233064P.  
PR 14-SEP-2000; 2000US-0233065P.  
PR 21-SEP-2000; 2000US-0234223P.  
PR 21-SEP-2000; 2000US-0234274P.  
PR 25-SEP-2000; 2000US-0234977P.  
PR 25-SEP-2000; 2000US-0234988P.  
PR 26-SEP-2000; 2000US-0235484P.  
PR 27-SEP-2000; 2000US-0235634P.  
PR 27-SEP-2000; 2000US-0235635P.  
PR 29-SEP-2000; 2000US-0236327P.  
PR 29-SEP-2000; 2000US-0236367P.  
PR 29-SEP-2000; 2000US-0236368P.  
PR 29-SEP-2000; 2000US-0236369P.  
PR 29-SEP-2000; 2000US-0236370P.  
PR 02-OCT-2000; 2000US-0236802P.  
PR 02-OCT-2000; 2000US-0237037P.  
PR 02-OCT-2000; 2000US-0237038P.  
PR 02-OCT-2000; 2000US-0237039P.  
PR 02-OCT-2000; 2000US-0237040P.  
PR 13-OCT-2000; 2000US-0239935P.  
PR 13-OCT-2000; 2000US-0239937P.  
PR 20-OCT-2000; 2000US-0240960P.  
PR 20-OCT-2000; 2000US-0241211P.  
PR 20-OCT-2000; 2000US-0241855P.  
PR 20-OCT-2000; 2000US-0241866P.  
PR 20-OCT-2000; 2000US-0241877P.  
PR 20-OCT-2000; 2000US-0241808P.  
PR 20-OCT-2000; 2000US-0241809P.  
PR 20-OCT-2000; 2000US-0241826P.  
PR 01-NOV-2000; 2000US-0244617P.  
PR 08-NOV-2000; 2000US-0246474P.  
PR 08-NOV-2000; 2000US-0246475P.  
PR 08-NOV-2000; 2000US-0246476P.  
PR 08-NOV-2000; 2000US-0246477P.  
PR 08-NOV-2000; 2000US-0246478P.  
PR 08-NOV-2000; 2000US-0246523P.  
PR 08-NOV-2000; 2000US-0246524P.  
PR 08-NOV-2000; 2000US-0246525P.  
PR 08-NOV-2000; 2000US-0246526P.  
PR 08-NOV-2000; 2000US-0246527P.  
PR 08-NOV-2000; 2000US-0246528P.  
PR 08-NOV-2000; 2000US-0246532P.  
PR 08-NOV-2000; 2000US-0246609P.

PR 08-NOV-2000; 2000US-0246610P.  
PR 08-NOV-2000; 2000US-0246611P.  
PR 08-NOV-2000; 2000US-0246613P.  
PR 17-NOV-2000; 2000US-0249207P.  
PR 17-NOV-2000; 2000US-0249208P.  
PR 17-NOV-2000; 2000US-0249209P.  
PR 17-NOV-2000; 2000US-0249210P.  
PR 17-NOV-2000; 2000US-0249211P.  
PR 17-NOV-2000; 2000US-0249212P.  
PR 17-NOV-2000; 2000US-0249213P.  
PR 17-NOV-2000; 2000US-0249214P.  
PR 17-NOV-2000; 2000US-0249215P.  
PR 17-NOV-2000; 2000US-0249216P.  
PR 17-NOV-2000; 2000US-0249217P.  
PR 17-NOV-2000; 2000US-0249218P.  
PR 17-NOV-2000; 2000US-0249244P.  
PR 17-NOV-2000; 2000US-0249245P.  
PR 17-NOV-2000; 2000US-0249246P.  
PR 17-NOV-2000; 2000US-0249265P.  
PR 17-NOV-2000; 2000US-0249297P.  
PR 17-NOV-2000; 2000US-0249299P.  
PR 17-NOV-2000; 2000US-0249300P.  
PR 01-DEC-2000; 2000US-0250160P.  
PR 01-DEC-2000; 2000US-0250391P.  
PR 05-DEC-2000; 2000US-0251030P.  
PR 05-DEC-2000; 2000US-0251988P.  
PR 05-DEC-2000; 2000US-0256719P.  
PR 06-DEC-2000; 2000US-0251479P.  
PR 08-DEC-2000; 2000US-0251856P.  
PR 08-DEC-2000; 2000US-0251868P.  
PR 08-DEC-2000; 2000US-0251869P.  
PR 08-DEC-2000; 2000US-0251899P.  
PR 08-DEC-2000; 2000US-0251900P.  
PR 11-DEC-2000; 2000US-0254907P.  
PR 05-JAN-2001; 2001US-0255678P.

## (HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-483426/52.

Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
useful for preventing, diagnosing and/or treating cancers and metastasis.  
Disclosure; SEQ ID NO 38234; 3071pp + Sequence Listing; English.

AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)  
CC amino acid sequences given in AAM62170 to AAM91921. (I) have cytosolic  
CC activity, and can be used in gene therapy and vaccine production. (I)  
CC proteins and polynucleotides may be used in the prevention, diagnosis and  
CC treatment of diseases associated with inappropriate (I) expression. For  
CC example, they may be used to treat disorders associated with decreased  
CC expression by rectifying mutations or deletions in a patient's genome  
CC that affect the activity of (I) by expressing inactive proteins or to  
CC supplement the patient's own production of (I). Additionally, (I)  
CC polynucleotides may be used to produce the secreted (I), by inserting the  
CC nucleic acids into a host cell and culturing the cell to express the  
CC protein. (I) proteins and polynucleotides may be used to prevent,  
CC diagnose and treat immune/haematopoietic-related diseases, especially  
CC cancers and cancer metastases of haematopoietic-derived cells. AAK64703  
CC to AAK67694 represent human immune/haematopoietic antigen genomic  
CC sequences from the present invention. AAK54942 to AAK54950 and AAM82169  
CC represent sequences used in the exemplification of the present invention  
XX  
XX  
SQ Sequence 1009 BP; 259 A; 269 C; 239 G; 242 T; 0 U; 0 Other;

Query Match 32.2%; Score 1005.8; DB 4; Length 1009;

Best Local Similarity 99.8%; Pred. No. 2.3e-180;

Matches 1007; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1766 CCGTCAGGCGCATTAACCGCTTAACGTCGTGAGAGGCTTTATTCCTTATTATAGAAACCGT 1825

Db 1009 CCGTACGGGCATTACCGCTAACGTCTGCAGAGCTTTATCCCTATTAAGAAAAACGCT 950  
Qy 1826 CACAGTACCTAGATCCCTCCGAGTTAATGATTAACATATGTCTGTTGGGGGCTCTT 1885  
Db 949 CACAGTACCTAGATCCCTCCGAGTTAATGATTAACATATGTCTGTTGGGGGCTCTT 890  
Qy 1886 TACAGGAGTCCGAGTGGTGGTCCGAGCCCTGCGAGCGTGGCCCCCTTCTGCGTGGGAC 1945  
Db 889 TACAGGAGTCCGAGTGGTGGTCCGAGCCCTGCGAGCGTGGCCCCCTTCTGCGTGGGAC 830  
Qy 1946 AGTTTGAAGAGTGGTGGTGGTGGAGTGAAGTTTGGAGAGGAGCGCTGTTTGTTCTATG 2005  
Db 829 AGTTTGAAGAGTGGTGGTGGTGGAGTGAAGTTTGGAGAGGAGCGCTGTTTGTTCTATG 770  
Qy 2006 TGGTGGTCTGTTTCCCGGACAAGAAAAATTGCATCAATGTCAAGAGCTTTTATTAAC 2065  
Db 769 TGGTGGTCTGTTTCCCGGACAAGAAAAATTGCATCAATGTCAAGAGCTTTTATTAAC 710  
Qy 2066 TTAATCTTTCAAGGCTTAATTTAGAGAGTGTCCGAGAGAGTTCATACAAAAGGCTT 2125  
Db 709 TTAATCTTTCAAGGCTTAATTTAGAGAGTGTCCGAGAGAGTTCATACAAAAGGCTT 650  
Qy 2126 TCTCTAAGACGGCTACAGCCCTTCTAGCAGAGTTTATCAATGTGCCCAAGAGCAGC 2185  
Db 649 TCTCTAAGACGGCTACAGCCCTTCTAGCAGAGTTTATCAATGTGCCCAAGAGCAGC 590  
Qy 2186 TGAAGAGATTTGAGGTCAATCTTCCACTGCGCTCAGGGGCTGACCTTATTAGAA 2245  
Db 589 TGAAGAGATTTGAGGTCAATCTTCCACTGCGCTCAGGGGCTGACCTTATTAGAA 530  
Qy 2246 ACCAAGAGAGTGGGTTGAACCTACTCTACGGAATTGATCAAGTGCACACTTGCT 2305  
Db 529 ACCAAGAGAGTGGGTTGAACCTACTCTACGGAATTGATCAAGTGCACACTTGCT 470  
Qy 2306 GCGAAGAGAGGCTCTCCCGACGACCCGAGATGGGGGTGAAGAGAGAGAGGCTTG 2365  
Db 469 GCGAAGAGAGGCTCTCCCGACGACCCGAGATGGGGGTGAAGAGAGAGAGGCTTG 410  
Qy 2366 GGGTGGGCTCACTGTGTTTAAACAGGACATTTCTCTCTGCGGCTTATTTTGT 2425  
Db 409 GGGTGGGCTCACTGTGTTTAAACAGGACATTTCTCTCTGCGGCTTATTTTGT 350  
Qy 2426 CAGAACTAGACAGAGTGTGTTGAACCTCTTTGACAGAGGCTGGGAATCTCTTAAAG 2485  
Db 349 CAGAACTAGACAGAGTGTGTTGAACCTCTTTGACAGAGGCTGGGAATCTCTTAAAG 290  
Qy 2486 CACTTAATCTTAATTTATCCCTGGAATGTGCGTCTGCGCAGTAGAGGGCTGGCTTGG 2545  
Db 289 CACTTAATCTTAATTTATCCCTGGAATGTGCGTCTGCGCAGTAGAGGGCTGGCTTGG 230  
Qy 2546 CAGCTTCCCTGACCCCGCGCTGCGCGCCCTCCGGGGTATGTGGCAATTACCTGGCCACA 2605  
Db 229 CAGCTTCCCTGACCCCGCGCTGCGCGCCCTCCGGGGTATGTGGCAATTACCTGGCCACA 170  
Qy 2606 GAGGTTTGAAGCAATCAGCTCTGAGACTGGGTTAGATGTAAACAGCTTTAATTGGAGT 2665  
Db 169 GAGGTTTGAAGCAATCAGCTCTGAGACTGGGTTAGATGTAAACAGCTTTAATTGGAGT 110  
Qy 2666 TTAAGAGCTTTTAAAAAGTAAATATCTCTGAAAGAAAAATGACGTAAACAAGGCTGT 2725  
Db 109 TTAAGAGCTTTTAAAAAGTAAATATCTCTGAAAGAAAAATGACGTAAACAAGGCTGT 50  
Qy 2726 ACTATGAAGCTGTATTTTATTAAGAACGCTGGGCCCTGAACCTATA 2774  
Db 49 ACTATGAAGCTGTATTTTATTAAGAACGCTGGGCCCTGAACCTATA 1

DT 06-NOV-2001 (first entry)  
XX Human immune/haematopoietic antigen encoding cDNA SEQ ID NO:7845.  
DE Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
KW cytostatic; gene therapy; vaccine; metastasis; ss.  
XX Homo sapiens.  
XX WO200157182-A2.  
XX 09-AUG-2001.  
PD 17-JAN-2001; 2001MO-US001354.  
PF 31-JAN-2000; 2000US-0179065P.  
PR 04-FEB-2000; 2000US-0180628P.  
PR 24-FEB-2000; 2000US-0184664P.  
PR 02-MAR-2000; 2000US-0186350P.  
PR 16-MAR-2000; 2000US-0189874P.  
PR 17-MAR-2000; 2000US-0190075P.  
PR 18-APR-2000; 2000US-0198123P.  
PR 19-MAY-2000; 2000US-0205515P.  
PR 07-JUN-2000; 2000US-0209467P.  
PR 28-JUN-2000; 2000US-0214886P.  
PR 30-JUN-2000; 2000US-0215135P.  
PR 07-JUL-2000; 2000US-0216647P.  
PR 07-JUL-2000; 2000US-0216880P.  
PR 11-JUL-2000; 2000US-0217487P.  
PR 14-JUL-2000; 2000US-0218290P.  
PR 26-JUL-2000; 2000US-0220963P.  
PR 26-JUL-2000; 2000US-0220964P.  
PR 14-AUG-2000; 2000US-0224518P.  
PR 14-AUG-2000; 2000US-0224519P.  
PR 14-AUG-2000; 2000US-0225213P.  
PR 14-AUG-2000; 2000US-0225214P.  
PR 14-AUG-2000; 2000US-0225266P.  
PR 14-AUG-2000; 2000US-0225267P.  
PR 14-AUG-2000; 2000US-0225268P.  
PR 14-AUG-2000; 2000US-0225270P.  
PR 14-AUG-2000; 2000US-0225477P.  
PR 14-AUG-2000; 2000US-0225757P.  
PR 14-AUG-2000; 2000US-0225758P.  
PR 14-AUG-2000; 2000US-0225759P.  
PR 18-AUG-2000; 2000US-0226279P.  
PR 22-AUG-2000; 2000US-0226811P.  
PR 22-AUG-2000; 2000US-0226868P.  
PR 22-AUG-2000; 2000US-0227182P.  
PR 23-AUG-2000; 2000US-0227009P.  
PR 30-AUG-2000; 2000US-0228924P.  
PR 01-SEP-2000; 2000US-0228287P.  
PR 01-SEP-2000; 2000US-0229343P.  
PR 01-SEP-2000; 2000US-0229344P.  
PR 01-SEP-2000; 2000US-0229345P.  
PR 05-SEP-2000; 2000US-0229509P.  
PR 05-SEP-2000; 2000US-0229513P.  
PR 06-SEP-2000; 2000US-0230437P.  
PR 06-SEP-2000; 2000US-0230438P.  
PR 08-SEP-2000; 2000US-0231242P.  
PR 08-SEP-2000; 2000US-0231243P.  
PR 08-SEP-2000; 2000US-0231244P.  
PR 08-SEP-2000; 2000US-0231413P.  
PR 08-SEP-2000; 2000US-0231414P.  
PR 08-SEP-2000; 2000US-0232080P.  
PR 08-SEP-2000; 2000US-0232081P.  
PR 12-SEP-2000; 2000US-0231968P.  
PR 14-SEP-2000; 2000US-0233397P.  
PR 14-SEP-2000; 2000US-0233398P.  
PR 14-SEP-2000; 2000US-0233399P.  
PR 14-SEP-2000; 2000US-0233400P.  
PR 14-SEP-2000; 2000US-0233401P.  
PR 14-SEP-2000; 2000US-0233063P.

RESULT 6  
AAK62785/c  
ID AAK62785 standard; cDNA; 973 BP.  
XX AAK62785;  
AC  
XX



PR 14-SEP-2000; 2000US-0233064P.  
 PR 14-SEP-2000; 2000US-0233065P.  
 PR 21-SEP-2000; 2000US-0234223P.  
 PR 21-SEP-2000; 2000US-0234274P.  
 PR 25-SEP-2000; 2000US-0234997P.  
 PR 25-SEP-2000; 2000US-0234998P.  
 PR 26-SEP-2000; 2000US-0235484P.  
 PR 27-SEP-2000; 2000US-0235834P.  
 PR 27-SEP-2000; 2000US-0235836P.  
 PR 29-SEP-2000; 2000US-0236327P.  
 PR 29-SEP-2000; 2000US-0236367P.  
 PR 29-SEP-2000; 2000US-0236368P.  
 PR 29-SEP-2000; 2000US-0236369P.  
 PR 29-SEP-2000; 2000US-0236370P.  
 PR 02-OCT-2000; 2000US-0236802P.  
 PR 02-OCT-2000; 2000US-0237037P.  
 PR 02-OCT-2000; 2000US-0237038P.  
 PR 02-OCT-2000; 2000US-0237039P.  
 PR 02-OCT-2000; 2000US-0237040P.  
 PR 13-OCT-2000; 2000US-0239935P.  
 PR 13-OCT-2000; 2000US-0239937P.  
 PR 20-OCT-2000; 2000US-0240960P.  
 PR 20-OCT-2000; 2000US-0241212P.  
 PR 20-OCT-2000; 2000US-0241785P.  
 PR 20-OCT-2000; 2000US-0241786P.  
 PR 20-OCT-2000; 2000US-0241787P.  
 PR 20-OCT-2000; 2000US-0241808P.  
 PR 20-OCT-2000; 2000US-0241809P.  
 PR 20-OCT-2000; 2000US-0241826P.  
 PR 01-NOV-2000; 2000US-0244617P.  
 PR 08-NOV-2000; 2000US-0246474P.  
 PR 08-NOV-2000; 2000US-0246475P.  
 PR 08-NOV-2000; 2000US-0246476P.  
 PR 08-NOV-2000; 2000US-0246477P.  
 PR 08-NOV-2000; 2000US-0246478P.  
 PR 08-NOV-2000; 2000US-0246523P.  
 PR 08-NOV-2000; 2000US-0246524P.  
 PR 08-NOV-2000; 2000US-0246525P.  
 PR 08-NOV-2000; 2000US-0246526P.  
 PR 08-NOV-2000; 2000US-0246527P.  
 PR 08-NOV-2000; 2000US-0246528P.  
 PR 08-NOV-2000; 2000US-0246532P.  
 PR 08-NOV-2000; 2000US-0246609P.  
 PR 08-NOV-2000; 2000US-0246610P.  
 PR 08-NOV-2000; 2000US-0246611P.  
 PR 08-NOV-2000; 2000US-0246613P.  
 PR 17-NOV-2000; 2000US-0249207P.  
 PR 17-NOV-2000; 2000US-0249208P.  
 PR 17-NOV-2000; 2000US-0249209P.  
 PR 17-NOV-2000; 2000US-0249210P.  
 PR 17-NOV-2000; 2000US-0249211P.  
 PR 17-NOV-2000; 2000US-0249212P.  
 PR 17-NOV-2000; 2000US-0249213P.  
 PR 17-NOV-2000; 2000US-0249214P.  
 PR 17-NOV-2000; 2000US-0249215P.  
 PR 17-NOV-2000; 2000US-0249216P.  
 PR 17-NOV-2000; 2000US-0249217P.  
 PR 17-NOV-2000; 2000US-0249218P.  
 PR 17-NOV-2000; 2000US-0249244P.  
 PR 17-NOV-2000; 2000US-0249245P.  
 PR 17-NOV-2000; 2000US-0249246P.  
 PR 17-NOV-2000; 2000US-0249265P.  
 PR 17-NOV-2000; 2000US-0249297P.  
 PR 17-NOV-2000; 2000US-0249299P.  
 PR 17-NOV-2000; 2000US-0249300P.  
 PR 01-DEC-2000; 2000US-0250316P.  
 PR 01-DEC-2000; 2000US-0250391P.  
 PR 05-DEC-2000; 2000US-0251030P.  
 PR 05-DEC-2000; 2000US-0251988P.  
 PR 05-DEC-2000; 2000US-0256719P.  
 PR 06-DEC-2000; 2000US-0251479P.  
 PR 08-DEC-2000; 2000US-0251856P.  
 PR 08-DEC-2000; 2000US-0251868P.

PR 08-DEC-2000; 2000US-0251869P.  
 PR 08-DEC-2000; 2000US-0251989P.  
 PR 08-DEC-2000; 2000US-0251990P.  
 PR 11-DEC-2000; 2000US-0254097P.  
 PR 05-JAN-2001; 2001US-0259678P.  
 XX  
 XX (HUMA-) HUMAN GENOME SCI INC.  
 PA  
 PI Rosen CA, Barash SC, Ruben SM;  
 XX WPI, 2001-483426/52.  
 DR P-PSDB; AAM90004.  
 XX  
 PT Nucleic acids encoding human immune/hematopoietic antigen polypeptides,  
 useful for preventing, diagnosing and/or treating cancers and metastasis.  
 XX  
 PS Claim 1; SEQ ID NO 7845; 3071pp + Sequence Listing; English.  
 XX  
 CC AAK54951 to AAK64702 encode the human immune/hematopoietic antigen (I)  
 CC amino acid sequences given in AAM62170 to AAM91921. (I) have cytostatic  
 CC activity, and can be used in gene therapy and vaccine production. (I)  
 CC proteins and polynucleotides may be used in the prevention, diagnosis and  
 CC treatment of diseases associated with inappropriate (I) expression. For  
 CC example, they may be used to treat disorders associated with decreased  
 CC expression by rectifying mutations or deletions in a patient's genome  
 CC that affect the activity of (I) by expressing inactive proteins or to  
 CC supplement the patient's own production of (I). Additionally, (I)  
 CC polynucleotides may be used to produce the secreted (I), by inserting the  
 CC nucleic acids into a host cell and culturing the cell to express the  
 CC protein. (I) proteins and polynucleotides may be used to prevent,  
 CC diagnose and treat immune/hematopoietic-related diseases, especially  
 CC cancers and cancer metastases of hematopoietic-derived cells. AAK64703  
 CC to AAK67694 represent human immune/hematopoietic antigen genomic  
 CC sequences from the present invention. AAK54942 to AAK54950 and AAM62169  
 CC represent sequences used in the exemplification of the present invention  
 XX  
 SQ Sequence 973 BP; 255 A; 253 C; 226 G; 234 T; 0 U; 5 Other;  
 Query Match 28.3%; Score 883.2; DB 4; Length 973;  
 Best Local Similarity 96.6%; Pred. No. 3.2e-157;  
 Matches 941; Conservative 1; Mismatches 28; Indels 4; Gaps 4;  
 QY 1803 ATTCCCTAATATAAACAACGTCACAGTACCTTAGATCCCTCGAGTTAATGAGTTAA 1862  
 DB 973 ATCCCTAATATAAACAACNTTCAAGTACCTTAGATCCCTCGAGTTAATGAGTTAA 914  
 QY 1863 CACATGTCGTGTGGGGGCTTTTACAGGAGTCCGAGTTCGGTCCCAACCCCTG-CCAG 1921  
 DB 913 ACATGTCGTGTGGGAGTCTTTACAGGAGTCCGAGTTCGGTCCCAACCCCTGCCAG 854  
 QY 1922 CGTGGCCCCCTTTCGCGTGGGACGTTTGAAGAAGTGGTGGGTGAGTGAAGTTGG 1981  
 DB 853 CGTGGCCCCCTTTCGCTGGGACGTTTGAAGAAGTGGTGGGTGAGTGAAGTTGG 794  
 QY 1992 AGAGGAGCGCTGTTGTTTATGATGTTGCTGTTTCCCGGACAAAGAAATTGCAAT 2041  
 DB 793 AGAGGAGCGCTGTTGTTTATGATGTTGCTGTTTCCCGGACAAAGAAATTGCAAT 734  
 QY 2042 CAAATGTACAGCTTTTATTACCTTAATCTTTCAGGGGCTTAATTAGAGAGTGTCTT 2101  
 DB 733 CAAATGTACAGCTTTTATTACCTTAATCTTTCAGGGGCTTAATTAGAGAGTGTCTT 674  
 QY 2102 GAGAGCACTTCAAAAGAGGCTTCTTAAGAGCGGCTACAGCCCTTCTTACAGAGTT 2161  
 DB 673 GAGAGCACTTCAAAAGAGGCTTCTTAAACCGGCTACAGCCCTTCTTACAGAGTT 614  
 QY 2162 TATCATTGTCCTCCCAAGAGCAGTAAAGAGATTGAGGTATGATCACTCCCACTGCCG 2221  
 DB 613 TATCATTGTCCTCCCAAGAGCAGTAAAGAGATTGAGGTATGATCACTCCCACTGCCG 554  
 QY 2222 TCAGGGGCTGACCTTAATTAGAGAAACAAAGAGGCTGGTGAACCTTCTCAGGACT 2281  
 DB 553 TCAGGGGCTGACCTTAATTAGAGAAACAAAGAGGCTGGTGAACCTTCTCAGGACT 494





Db 161 GGGGGGGGGGCTGAGATTTTGTATACGTTAGCGTCGATTTTTCGTCGTGTTTG 102  
 QY 1455 CAGGAGCGGGGGGGGGTTCGACCCAGAGAGCCCTGCGCCATCTTTTCGGCGCC 1514  
 Db 101 TAGGAGCGCGGGGGGGTTCGATTTTAGAGAGTTTGTGTTATTTTTCGGCGTC 42  
 QY 1515 GGGCTGCTGGGCGGCTGTGGCCCTAGCCCGTGTGCGTGGCGAA 1555  
 Db 41 GTGTGTGGCGGTTGTGTGTTTATGTCGTGCGTGGCGAA 1

## RESULT 9

ABQ40780  
 ID ABQ40780 standard; DNA; 1061 BP.

XX ABQ40780;  
 XX 12-JUL-2002 (first entry)

XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 27371.

XX Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
 KW drug; side effect; cancer; central nervous system; cardiovascular;  
 KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
 KW SNP; cell differentiation; ds.

XX Homo sapiens.

XX WO200218632-A2.

XX 07-MAR-2002.

XX 01-SEP-2001; 2001WO-EP010074.

XX 01-SEP-2000; 2000DE-01043826.

XX 05-SEP-2000; 2000DE-01044543.

XX (EPIC-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K, Gnetig D;  
 WI; 2002-371829/40.

XX Determining the degree of cytosine methylation in genomic DNA, useful for  
 PT diagnosis and prognosis, comprises selective hybridization of amplicons  
 PT from chemically treated DNA.

XX Claim 12; 56pp + sequence listing; 56pp; German.

XX This invention describes a novel method for determining the degree of  
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
 CC genomic sample of DNA. The sample is treated chemically to convert  
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
 CC DNA that contains the target C is amplified to form a labeled amplicon.  
 CC The amplicon is hybridized to two classes, each with at least one member,  
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the  
 CC degree of hybridization to both classes is determined from the label on  
 CC the amplicon. From the ratio of labels hybridized to the two classes of  
 CC oligomers, the degree of methylation is calculated. The method is used:  
 CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs  
 CC and of a wide range of diseases, e.g. cancer, disorders of the central  
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,  
 CC particularly by detecting mutations or single nucleotide polymorphisms  
 CC (SNP's); and (ii) for differentiation of cell or tissue types and for  
 CC investigating cell differentiation. The method allows the methylation  
 CC status of many C residues to be determined simultaneously. ABQ1410-  
 CC ABQ4112 represent genomic DNA sequences used to illustrate the method  
 CC for determining the degree of cytosine methylation described in the  
 CC disclosure of the invention

XX Sequence 1061 BP; 147 A; 137 C; 429 G; 348 T; 0 U; 0 Other;

XX Query Match 23.3%; Score 726.6; DB 6; Length 1061;

Best Local Similarity 80.3%; Pred. No. 1-2e-127;  
 Matches 852; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

QY 495 GCTTCCCGGAGATTGGACCCACGAGAGATGGGAGCCGACCTCAGCTTTCGAGGAG 554  
 Db 1 GTTGTGTTTCGAGATTGATTTACGAGAGATGGGAGATGTTTATTTTATGTTTCGTAAGGAG 60  
 QY 555 CCACCGTGGAGAGCCAGGCGGTGACAGACACGACGTGTGACTGGAGTGGCGCTGGGGA 614  
 Db 61 TTATCGTGGAGATTAGGCGGAGTGTAGAGATACGACGTGTGATTCGAGATGCGTTTGGGGA 120  
 QY 615 GAATGACGAGAGGAGCGGGGAGCCGCTAACGAGGAGCTCCCTCTGGCGCGCCCGTCCGACGA 674  
 Db 121 GGATGACGAGAGGAGCGGGGAGATCGTTAACGGGGTTTTCGCGCTTTCGTTGATGAGA 180  
 QY 675 GCGGACGTCGAGGAGTCCCGGGGCGGCTCCGTGACGTTGGCGGTGAGCGGACGAGTC 734  
 Db 181 GCGGTACGTCGAGGAGTTCGGGCGGAGTTTCGTGACGTTTGGCGGTGAGCGGAGTT 240  
 QY 735 ACGACCAATGAGAGCGCTTGTGCGCGGCGGCCACAGGCGGGAGTGGAGTACCAT 794  
 Db 241 ACGATTATGAGAGCGCTTGTGTCGCGGATTTAAGGTGAGAGTGGAGTGAATAT 300  
 QY 795 CCGCGCGCTGAGGAGGAGCTTAACGAGCGCGGCGCGCGCCAGCCGAGCCGAC 854  
 Db 301 TTTGTCGCTTGAAGGAGGAGTTAACGAGCGCGGCGGCTCGGGTTTATGCGAGTTATC 360  
 QY 855 GCGATGCGAGAGAGAGAGTGCACAGCGCTGCTGACGCGGCTCAACAGACGATCGCTGC 914  
 Db 361 GCGATGCGAGAGAGAGTGTAAAGCGCTTGTGACGCGGTTTAAGACATATGCGCTGT 420  
 QY 915 TACCAACCACTGATGCTGACCGCTGAGTGCCTCGCGGACCTGCGAGAACCTGCGGACAGAG 974  
 Db 421 TATTATTATTTTGTGTGATGCTGAGTGTGCGGAGTTGTAGAAATTTGCGGTAGAG 480  
 QY 975 CTGCAAAAGACGCGCCAGAGAGCGCAGAGAGCTGCGGTGCTCAACTGCGCCGCTGACT 1034  
 Db 481 TTGTAAAGAGCGCTTGAAGAGCGGTAGAGAGTGGCGGTTTATTTGGCTTCGTTGAT 540  
 QY 1035 GCTGTGCTGCGGACCGGAGCGCTGCGCGGACGACGAGCGCGGAGTTGAGCGGCTCTGG 1094  
 Db 541 GTTGTGTTCCGAGATCGGAGGTTTGTGCTGTCAACAGCGCGCTGAGTTGACCGGTTTGG 600  
 QY 1095 GTGGCTTCTGGGCGTGCCTGACCTGTGTAAGCGGACATGCGAGCGCTGCTGAGAGCTG 1154  
 Db 601 GTGTTTTCGAGGTTGTTGTTGATTTGTTGAAAGCGAATATGCAACGCGCGTTGAGTTG 660  
 QY 1155 GCGCGCGGTTCCGCTGACAGCGCGCGCGGACCGCTGTGTCGACAGGTGTGGCTGGC 1214  
 Db 661 GCGGTGCGGTTTGTGTTGTAACGCGCGCGGCTGTGTTGCGTAATAGTGTGGTTGGC 720  
 QY 1215 GCTCTCTCGCGCGTGGCGCGCGCGCGCTGAGACCGCGACCTGCGGCTGAGAGCGGAG 1274  
 Db 721 GTTTCCTTCGCGCGTGGCGCGCGCGCGCTGTGAGTTCCTAATTTTGGCGTTGAGAGCGGAG 780  
 QY 1275 GCGGACTTGAAGTTCGCGGAGCTGCGGAGGCTGAGCGCGAGGTCCTTCAGGTGGGCGAG 1334  
 Db 781 GCGGATTCGACGTCGCGAATTTGCGGAGATTGAGCGCGAGGTTTTCAGTGGGCGAG 840  
 QY 1335 ATGATGCAACAATGAGAGTGAAGTCAAGTGCCTGCGGCTGAGACGTCGAGACCGCGAG 1394  
 Db 841 ATGATGCAATAATGAGAGTGAAGTCAAGTGTGTTGTTGATGCTGTAAGTTCCGATG 900  
 QY 1395 GCGGCGGCGCGGAGCTCTGTCAACGCTGACGCGCGGCGCTCTCGTGTGTGTCTTG 1454  
 Db 901 GCGGCGGCGCGTCAAGTTTGTTCACGTTGACGTTGCGTTCGTTTTCGTCGTGTTTGG 960  
 QY 1455 CAGGAGCGCGGGGGGGTGGAGCCCGAGAGGCGCGGCGCATCTTTTCGGCGCC 1514  
 Db 961 TAGGAGCGCGGGGGGGTGGAGTTTGAAGAGTTTGTGTCGTTATTTTTCGGCGTC 1020  
 QY 1515 GTGCTGTGCGCGCTGTGTGCTTACCGCTGTGTGCGTGGCGAA 1555

DB 1021 GTGTTGTGGCGGTGTGTTTACGTGCTGTCGTCGCA 1061

RESULT 10

ID ACH87504/c

ACH87504 standard; DNA; 708 BP.

XX ACH87504;

DT 29-JUL-2004 (first entry)

XX Human genome derived single exon probe #20699.

DE Human genome derived single exon probe; microarray;

KW Human; probe; ss; gene expression; single exon probe; microarray;

XX alternative splicing event; genomic alteration.

OS Homo sapiens.

XX US2003194704-A1.

XX 16-OCT-2003.

XX 03-APR-2002; 2002US-00029386.

XX 03-APR-2002; 2002US-00029386.

XX (PENN/) PENN S G.

PA (RANK/) RANK D R.

PA (HANZ/) HANZEL D K.

XX Penn SG, Rank DR, Hanzel DK;

XX MPI; 2004-119264/12.

DR New human genome-derived single exon nucleic acid probes useful for human

PT gene expression analysis, for identifying or characterizing alternative

PT splicing events, for assessing genomic alterations or as tools for

PT surveying tissues.

XX Claim 1; SEQ ID NO 20699; 80pp; English.

XX The invention relates to a nucleic acid probe for measuring human gene

XX expression, comprising any of the 27,400 fully defined nucleotide

XX sequences in the specification, or their complements or fragments, and

XX encoding at least 8 amino acids of any of the 688 amino acid sequences

XX fully defined in the specification. The probe is a single exon probe that

XX hybridizes under high stringency conditions to a nucleic acid molecule

XX expressed in human cells or tissues. Also included are a spatially-

XX addressable set of single exon nucleic acid probes for measuring human

XX gene expression (comprising a plurality of single exon nucleic acid

XX probes cited above, where each of the plurality of probes is separately

XX and addressably isolatable or amplifiable from the plurality), a single

XX exon microarray for measuring human gene expression, a method of

XX measuring human gene expression, a vector comprising the single exon

XX probe cited above, an ORF-encoded peptide comprising at least 8

XX contiguous amino acids of any of the above-mentioned amino acid

XX sequences (optionally with conservative amino acid substitutions), an

XX isolated antibody that binds specifically to a peptide cited above,

XX methods of selling and/or licensing single exon probes or microarrays to

XX a customer desiring to measure gene expression, a method of providing

XX human gene expression data by subscription, and a computer-readable

XX storage medium which contains a database having a plurality of records

XX (each record including data on the expression of a single exon probe

XX cited above. The probe, methods and apparatus are useful in gene

XX expression analysis. The probes may be used as tools for surveying

XX tissues to detect the presence of expressed messages that contain their

XX specific exon, or in constructing genome-derived single exon microarrays.

XX In addition, the probes are used in identifying and characterizing

XX alternative splicing events, in detecting and characterizing gross

XX alterations in the genomic locus that includes their exon, in assessing

XX smaller genomic alterations, in priming the synthesis of nucleic acids,

XX or in expressing the ORF-encoded peptide. The present sequence is a human

XX single exon probe of the invention. Note: The sequence data for this

CC patent did not form part of the printed specification, but was obtained

CC in electronic format directly from USPTO at

CC seqdata.uspto.gov/sequence.html?docid=20030194704

XX

XX SQ Sequence 708 BP; 104 A; 279 C; 231 G; 94 T; 0 U; 0 Other;

XX

XX Query Match 22.6%; Score 706.4; DB 12; Length 708;

XX Best Local Similarity 99.9%; Pred. No. 7.4e-124;

XX Matches 707; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 858 ATGGCAGAGGAGAGTCAAGGCGCTGTGACAGCGGCTCAACAACAGCTGGTCTAC 917

DB 708 ATGGCAGAGGAGAGTCAAGGCGCTGTGACAGCGGCTCAACAACAGCTGGTCTAC 649

QY 918 CACCACTGTGTGACCGTGGTGTGCTGGGGAATCGACAGACTGGGGCAGAGCTG 977

DB 648 CACCACTGTGTGACCGTGGTGTGCTGGGGAATCGACAGACTGGGGCAGAGCTG 589

QY 978 CAAAAGACCGCAGAGAGCGAGAGCTGGGCTGTGACAGCTGGCCCGGCTGACTGCT 1037

DB 588 CAAAAGACCGCAGAGAGCGAGAGCTGGGCTGTGACAGCTGGCCCGGCTGACTGCT 529

QY 1038 GTGCTGCGGACCGGGGCTGTGCGCGCGACAGAGCGCGGAGTTGAGCGGCTTGGGTTG 1097

DB 528 GTGCTGCGGACCGGGGCTGTGCGCGCGACAGAGCGCGGAGTTGAGCGGCTTGGGTTG 469

QY 1098 GCCTTCTGGGCTGTGCTGTGACCTGTGTGAAAGGAGACATCGCTCCGTGAGCTGGGC 1157

DB 468 GCCTTCTGGGCTGTGCTGTGACCTGTGTGAAAGGAGACATCGCTCCGTGAGCTGGGC 409

QY 1158 GCGGCGTCCCGCTGACACGCGCGGCGAGCGCTGTGTCGACAGAGTGTGCTGGCGCC 1217

DB 408 GCGGCGTCCCGCTGACACGCGCGGCGAGCGCTGTGTCGACAGAGTGTGCTGGCGCC 349

QY 1218 TCTTCGCGGTGCGCGCGCGCGCTGTGACACCTCGGCTTGGCTGTGAGCGGAGGCT 1277

DB 348 TCTTCGCGGTGCGCGCGCGCGCTGTGACACCTCGGCTTGGCTGTGAGCGGAGGCT 289

QY 1278 GACTTGACGTGCGGACCTGTGCGGAGCTGTGAGCGGAGGCTTTCAGGTGGCGGAGTG 1337

DB 288 GACTTGACGTGCGGACCTGTGCGGAGCTGTGAGCGGAGGCTTTCAGGTGGCGGAGTG 229

QY 1338 ATCGACAATGAGATGAGTCAACGTGCGCGCGCTGTGACCGTGCACCGCGCAGGCG 1397

DB 228 ATCGACAATGAGATGAGTCAACGTGCGCGCGCTGTGACCGTGCACCGCGCAGGCG 169

QY 1398 GCGGCGCGGAGCTGTGCAACGTGACGCGCGGCGCTTCTCGGTGTGCTTTCAG 1457

DB 168 GCGGCGCGGAGCTGTGCAACGTGACGCGCGGCGCTTCTCGGTGTGCTTTCAG 109

QY 1458 GAGCGCGGGGGGGTTCGACCCCGAGAGAGCGCTTGGCCGCACTTTTGGCGCGCTG 1517

DB 108 GAGCGCGGGGGGGTTCGACCCCGAGAGAGCGCTTGGCCGCACTTTTGGCGCGCTG 49

QY 1518 CTGCTGCGCGGCTGTGAGCCCTTACCGTGTGCTGTGCGAAGCTGAGCTGA 1565

DB 48 CTGCTGCGCGGCTGTGAGCCCTTACCGTGTGCTGTGCGAAGCTGAGCTGA 1

RESULT 11

ABQ40778/c

ID ABQ40778 standard; DNA; 1061 BP.

XX ABQ40778;

XX 12-JUL-2002 (first entry)

XX Oligonucleotide for detecting cytosine methylation SEQ ID NO 27369.

DE Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; digoxins;

KW drug; side effect; cancer; central nervous system; cardiovascular;

KW gastrointestinal; respiratory system; single nucleotide polymorphism;

KW SNP; cell differentiation; ds.





XX 01-SBP-2000; 2000DE-01043826.  
 PR 05-SBP-2000; 2000DE-01044543.  
 XX (EPiG-) EPIGENOMICS AG.  
 PA Olek A, Piepenbrock C, Berlin K, Gnetig D;  
 XX WPI; 2002-371829/40.  
 DR  
 XX  
 PT Determining the degree of cytosine methylation in genomic DNA, useful for  
 PT diagnosis and prognosis, comprises selective hybridization of amplicons  
 PT from chemically treated DNA.  
 XX  
 PS Claim 12; 56pp + Sequence Listing; 56pp; German.  
 CC This invention describes a novel method for determining the degree of  
 CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
 CC genomic sample of DNA. The sample is treated chemically to convert  
 CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
 CC DNA that contains the target C is amplified to form a labeled amplicon.  
 CC The amplicon is hybridised to two classes, each with at least one member,  
 CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the  
 CC degree of hybridisation to both classes is determined from the label on  
 CC the amplicon. From the ratio of labels hybridised to the two classes of  
 CC oligomers, the degree of methylation is calculated. The method is used:  
 CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs  
 CC and of a wide range of diseases, e.g. cancer, disorders of the central  
 CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,  
 CC particularly by detecting mutations or single nucleotide polymorphisms  
 CC (SNP/s); and (ii) for differentiation of cell or tissue types and for  
 CC investigating cell differentiation. The method allows the methylation  
 CC status of many C residues to be determined simultaneously. ABQ13410-  
 CC ABQ54121 represent genomic DNA sequences used to illustrate the method  
 CC for determining the degree of cytosine methylation described in the  
 CC disclosure of the invention  
 CC  
 SQ Sequence 1061 BP; 439 A; 344 C; 137 G; 141 T; 0 U; 0 Other;  
 Query Match 19.0%; Score 594.4; DB 6; Length 1061;  
 Best Local Similarity 72.5%; Pred. No. 1.1e-102;  
 Matches 769; Conservative 0; Mismatches 291; Indels 0; Gaps 0;

QY 496 CTTCGCCCCGAGTTGGACCCACGAGATGGGGACCGGACCTTCGAGGGAGC 555  
 DB 2 CTTCACCCGAAATTAAACCCACGAAATAAATAAACCCTTCACTTCGCAAAAAAC 61  
 QY 556 CACCGTGAAGGCGAGGCGGTGCAAGACGAGTGTGCTCGAGTGGCGCTGGGAG 615  
 DB 62 CACCGTAAACCAAAACGATACAAACACGACGATATCTGAAATAGCGCTAAAAAA 121  
 QY 616 GATGACGAGGAGCGGGGACCGCTTAACGGGGCTCCCTGCGGCGCCCGTCCGACAG 675  
 DB 122 AATTAACGAAAAACCAAAAAACCGCTAACGAAATCTCCCTACGCGCCCGCTCGAAAA 181  
 QY 676 GCGCAGTGAAGGCTCCGCGGCGGCGCTCGTGAAGTTGCGCGTACGCGAGCGAGTCA 735  
 DB 182 AGCGACGTGAATTCGAAACGAACTCCGTAAACCTTAACGATACGCGCAACGATCA 241  
 QY 736 CGGACCATGAAGCGTTTCGCGCGCGCCCAAGCGCGGAGTGGGGTTAGCCACATC 795  
 DB 242 CGAACCATTAATAAGTTGTAACGCGGACCAAAACCGAATATAAATAATTAACGATC 301  
 QY 796 CTGCGCGGTGAGGGGAGGCTTAACGGGGCGCGGCGCGGCGCCAGCGGACCCACCG 855  
 DB 302 CTACCGCGCTAATAAAAAAACTTAACGAAACGAAACCGAACCCCAACCCACCG 361  
 QY 856 CGATGGCGAGGAGAGTCAAGCGCTGCTGGAAGGGCTCAACAAGACGATGCGTGT 915  
 DB 362 CGATTAAGAAAAAAATACAAACGCTACTTAACGAACTCAACAAACGACTACTACT 421  
 QY 916 ACCACCACTGTGCTGACCGTGGTGGCTCGCGGAGCTCGAGAACTTCGCGGACGAGC 975

DB 422 ACCACCACTTAATACTAACCGTGAATCTGACGAACTCGAAAACTTACGACAAAAAC 481  
 QY 976 TGCAAAAGACGCGCCAGAAAGGCGACGAGCTGGCGGTGTCACCTGGCGCCGGCTGACTG 1035  
 DB 482 TACAAAAACGCGCAAAAAACGAAATACTAACGATATCCACTTACGCGCCGACTTAACTA 541  
 QY 1036 CTGTGCTGCGGACCGGGGCTGGCGCGCGGACGAGGCGCGGAGTTGAGCGGCTTGGG 1095  
 DB 542 CTATCTACGCGACCGAAACCTTAACCGCGGACGAAACGCGCGCAATTCGAACGACTTAA 601  
 QY 1096 TGGCTTTCTGGGCTGCTGGAACCTGCTGGAAGCGGACATGCGAGCTGCTGAGACTGG 1155  
 DB 602 TAACCTTCGAACTTACCTTAACTTAAACGAACTAACGACGCGGCTTAAACTTA 661  
 QY 1156 GGGCGCGGTTCGCGTGAACGCGCGCGCGGCGCGCTGCTGAGGCTGAGGCTGGGG 1215  
 DB 662 AGCGCGCGTTCGCGTACAGCGCGCGGACGACCGCTTAATACGAAATTAATTAACG 721  
 QY 1216 CTTCTCCGCGCTGGCGCGCGCGCTGAGGACCGGACCTGCGGCTGAGGCGGAGG 1275  
 DB 722 CTTCTCCGCGCTTAACGAGCGCGCGGCTTAAACCGGCACTTACGACTCGAAACGAAAA 781  
 QY 1276 GCGACTTGAAGCTGCGGACCTGCGGAGCTGGAAGCGGAGGTTCTTCAAGTGGCGAGA 1335  
 DB 782 AGACTTGAAGCTGCGGACCTTAACTTAAACGCGAAATCCCTTCAATTAACGAAA 841  
 QY 1336 TGATGACAACTGAAGATGAAGTCAACGTGCGCGCGCTGAGCGGTCGACGCGCGGAG 1395  
 DB 842 TAAATGACAACTTAATAAATAAATAAATCAACGTACCGCGTAAACCGTACAAACCGACAAA 901  
 QY 1396 CGCGCGGCGCGCGAGCTCTGTCACAGTCAAGCGCGCGCGCTCTGCTGCTGCTTGC 1455  
 DB 902 CGACGACGCGGAACTCTGATCAAGATCAAGATCAAGCGGACCGCTCTGATCGATCTTAC 961  
 QY 1456 AGAGCGCGGGGGGAGTTCGACCCCGAGAGAGCCCTGAGCGGCAATCTTTTGGGCGCG 1515  
 DB 962 AAAAAACGGAATAAATTAACGACCCCAAAAAACCTTAACCGCATCTTTTGAACGCGG 1021  
 QY 1516 TGCTGCTGCGCGCTGCGCGCTGAGCGCGTGAAGCGGTGGCGGA 1555  
 DB 1022 TACTACTAACGACTTAACCTTAACCGGTATACGTAACGAA 1061

RESULT 13  
 ID ACH73793/C  
 ID ACH73793 standard; DNA; 524 BP.  
 XX ACH73793;  
 XX  
 AC 29-JUL-2004 (first entry)  
 XX  
 DT  
 XX  
 DB Human genome derived single exon probe #698.  
 XX  
 KW Human; probe; ss; gene expression; single exon probe; microarray;  
 KW alternative splicing event; genomic alteration.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US2003194704-A1.  
 PD 16-OCT-2003.  
 XX  
 PD 03-APR-2002; 2002US-00029386.  
 PF 03-APR-2002; 2002US-00029386.  
 PR  
 XX 03-APR-2002; 2002US-00029386.  
 XX  
 PA (PENN/) PENN S G.  
 PA (RANK/) RANK D R.  
 PA (HANZ/) HANZEL D K.  
 XX  
 PI Penn SG, Rank DR, Hanzel DK;  
 DR WPI; 2004-119264/12.











**THIS PAGE BLANK (USPTO)**



TITLE Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

JOURNAL Nature 420, 563-573 (2002)

REFERENCE 6 (bases 1 to 3429)

Aachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirokane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnate, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tanagi, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Retina RNA was provided by Dr. Stefano Gietz (Department of Neurobiology, Harvard Medical School, 220 Longwood Ave., Boston, MA 02115, USA) whose assistance is gratefully acknowledged. Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/

Location/Qualifiers

1. 3429

/organism="Mus musculus"

/mol\_type="mRNA"

/strain="C57BL/6J"

/db\_xref="PANTOM DB:A930005K04"

/db\_xref="taxon:10090"

/clone="A930005K04"

/tissue\_type="retina"

/clone\_id="RIKEN full-length enriched mouse cDNA library"

/dev\_stage="adult"

1. 3429

/note="unknown EST (GB|B170314, evidence: BLASTN, 99%, match=315)"

misc\_feature

ORIGIN

Query Match 22.4%; Score 700.6; DB 4; Length 3429;

Best Local Similarity 65.5%; Pred. No. 9e-112;

Matches 1440; Conservative 0; Mismatches 619; Indels 138; Gaps 23;

505 GAGTTGGACACCCAGAGATGGGAGACCTTCAGCTTCGAGGAGGACCGTGA 564

|||||

43 GACTTGGGATGATCAAGATGGGATCTTCACTTGAGGCTTCTCAGGAGCGCGTGAA 102

|||||

565 GGGCAGGCGGTGACAGACGACGCTGACCTGGAGTGGCTTGGGAGAGATGACGA 624

|||||

103 AAAAGAAAGTGTGAAAGACACTACGCTGACCGGAGCGGCTGCGGGGAAATGAGGA 162

|||||

625 GGGAGCGGGGAGCGCTAACGGGGCTCCCTGGGCGCGCGCTCGGAGAGGCGCAATC 684

|||||

163 GCGGATGGGAGGAAGACTAACACTACCCAGC-CGTGGGAGAGCGGAGGCGCACTTC 221

|||||

685 GAGGATCCCGGCGGAGCTCCCTGAGAGCTTGGCGGTAGCGCGGAGTCAAGTCAAGCATG 744

|||||

222 GAGGATCCCTGAGAGCTCTGTGAAACAGCGG-----CGAGAGCATG 265

|||||

745 AAGAGCTTCTGCGCGGCGCGCAAGCGCGGAGTGGGGATTAGCCACATCTCCCGGC 804

|||||

Db 266 AGAGCTTTCAGTCCGCGCTGCTCAAGGCTGAGATCAGGGCTGAGGCGCACTACAGCT 325

Qy 805 TGAAGGAGAGGCTTAACGAGGCGCGGCGCGGCGCGGCGGAGCCACCGCATGAGCA 864

Db 326 TGAAG-----GTTGCAAGTGGGATACCGCGGAGGAGCGGCGCATGGCCA 370

Qy 865 GGAAGAGTGCAGAGGCGCTGCTGAGACGGGCTCAACAGAGACTGGTGTACCAACACC 924

Db 371 GAAGAGAGTGCAGAGGCGCTGCTGAGACGGGCTCAACAGAGACCAAGGCTGTACATCACT 430

Qy 925 TGGTGTGACCGCGGAGGCTGGGCGGAGACTCGAGAACCTGGCGGCGAGAGCTCAAAAAG 984

Db 431 TAGTGTACTTGTGGGCGGCTCCGCGGACACTAGAGACTTACCGGAGAGCTCAGAAAG 490

Qy 985 CGGCGCAAGAGGCGCAGAGAGCTGGCGGATGCACTTGGCGCGGCGGCTGAGCTGTGTGC 1044

Db 491 CGGCGCAAGAGGCGGCGGAGAGCTGGCGGATGCACTTGGCGCGGCGGCTGAGCTGTGTGC 550

Qy 1045 GCG-ACCGGAGCTGGCGCGCGCGCGAGCGGCGGAGTTCAGAGCGCTTGGGTTGGCTTC 1103

Db 551 GGGCACCGGAGCTTACCTACCGAGAGCGCGCGGACTTGTAGCGGCTATGGGTGGCTTT 610

Qy 1104 TCGGGCTGCTGGAACCTGCTGGAAGGGAATGCGAGCGCTCGCTGAGAGCTGGGCGCGG 1163

Db 611 TCAAGCTGTGTGACCTGTCTGAAACCGAGATCAGCGGCACTTACGCTTGGAGCCACC 670

Qy 1164 TTCCCGCTGACCGCGCGCGCGCGAGCGCTGTGTGTGCGACAGTGTGCTGTGCGCTCTCC 1223

Db 671 CTTCGCTGTGACCGCGCGCGCGCGGCGGCACTTGTGTGCACTGGGGTGAACGCGGAT-CTAT 729

Qy 1224 GCGGTGGCGCGCGCGCGCGCTGAGACACCGGAGCTTGTGTGTGTGAGCGGAGGCGACTTC 1283

Db 730 GCTGTAGCTGCGCGCGCGCGCGGAGCGCGCGCGGCTGTGTGTGTGAGCGGAGCGACTTC 789

Qy 1284 GAGTGTGCGGACCTGTGGGAGGCTGTGAGCGGAGGAGTCTTCAAGTGGGCGAGATGATGAC 1343

Db 790 GATGTGCGGATGTGCGCGGAGTGTGAGCGGAGGAGTCTTCAAGTGGGCGAGATGATGAC 849

Qy 1344 AACATGAGATGAAAGTCAACGTGCGCGCGCTGTGACCGGTGACCGCGCGCGGCGGCG 1403

Db 850 GACATGAAATGAAAGTCAACGTGCGCGCGCTGTGACCGGTGACCGCGCGGCGGCGGAC 909

Qy 1404 GCGGAGCTCTGTGCAACGATCAAGCGCGCGCGCTCTCTCGGTG-----TGTCTTGCAG 1457

Db 910 GCCGAACTCTGTGCGGATGCAAGTGTGCGGCGCGGCGGCTGTGAGGCACTTGTGTGAG 969

Qy 1458 GAGGCGGCGGCGGAGTGTGCAACCGCGGAGGAGCGCGGCGGCGGCGGCGGCGGCGG 1517

Db 970 GAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1029

Qy 1518 CTGCTGCGGCTGTGCGGCTTACCGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1577

Db 1030 CTGT 1089

Qy 1578 GCGCGCTGCTGTGCGGCTTCTCTCTGAGAAAGACTTGGGAGTGGGTGTGTGTGTGTGT 1637

Db 1090 GGT 1136

Qy 1638 GCTGT 1696

Db 1137 AACCGTT-AAGAGAGAGAGT 1195

Qy 1697 AGTGCATGT 1751

Db 1196 AGCCGACCTGT 1255

Qy 1752 TAACTTTGCGCGCGCTGAGGCGATTAACGCTTACGCTTGTGTGTGTGTGTGTGTGTGT 1811

Db 1256 TAAACAGAC-AAGGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1314

Qy 1812 TAATAGAAACGCT-CACATGTACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1870

Db 1315 TAATAGAAATCGGTCAAGT 1355



QY 463 CGCTTAAGCGAGCGCGCTCTGACGCTTCTCCCGAGATTGCAACCAACGAG 522  
 Db 427 CGCTTAAGCGAGCGCGCTCTGACGCTTCTCCCGAGATTGCAACCAACGAG 486  
 QY 523 GATGGGACCGGACCGCTTCTGAGGAGGACCACTGTGAGGAGCGGTGCGAG 582  
 Db 487 GATGGGACCGGACCGCTTCTGAGGAGGACCACTGTGAGGAGCGGTGCGAG 546  
 QY 583 ACAAGAGTGTGACTCGGAGTGTGCGGAGGATGAGAGGAGGAGCGGAGCGCTA 642  
 Db 547 ACAAGAGTGTGACTCGGAGTGTGCGGAGGATGAGAGGAGGAGCGGAGCGCTA 606  
 QY 643 ACCGAGCTCCCTCTGCGCGCCGCTCGCAGAGGCGACGTCAGAGGTCCCGGCGCT 702  
 Db 607 ACCGAGCTCCCTCTGCGCGCCGCTCGCAGAGGCGACGTCAGAGGTCCCGGCGCT 666  
 QY 703 CCGTGAACGTT 713  
 Db 667 CCGTGAACGTT 677

RESULT 3  
 CFI45408 670 bp mRNA linear EST 06-AUG-2003  
 LOCUS UI-HF-CBO-aa0-e-10-0-UI.r1 NIH\_MGC\_210 Homo sapiens cDNA clone  
 ACCESSION CFI45408  
 VERSION CFI45408.1 GI:33260852  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Bakayote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homiidae; Homo.  
 1 (bases 1 to 670)  
 Bonaldo,M.F., Lennon,G. and Soares,M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)  
 JOURNAL PUBMED 8889548  
 COMMENT Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 315 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@iowa.edu  
 Tissue Procurement: Tim Ratliff  
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at  
 http://genome.uiowa.edu/distribution/humanfl.html  
 The following repetitive elements were found in this cDNA  
 sequence: 129-236, >MIR#SINE/MIR  
 Seq primer: PYX-5.

## FEATURES

source

location/Qualifiers

1..670  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAG:30569145"  
 /tissue\_type="CNCAP(3)T-225 cell line"  
 /lab\_host="DH10B (TI phage resistant)"  
 /clone\_id="NIH\_MGC\_210"  
 /note="Organ: Prostate; Vector: pT73 Pac; Site 1: Ecor I;  
 Site 2: Not I; The library was constructed according  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction, ligated

## ORIGIN

Query Match 21.2%; Score 662.4; DB 6; Length 670;  
 Best Local Similarity 99.8%; Pred. No. 5.2e-105;  
 Matches 663; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

with Ecor I adaptor, digested with Not I and then cloned  
 directionally into pT73 Pac vector. The library tag  
 sequence located between the Not I site and the polyA tail  
 is CCAC. Tissue was provided by Tim Ratliff."

QY 43 GCGCAACGCGGAGGAGCACTGACCCCGGCGGCGCCAGCCCTCGATATGCAATGAC 102  
 Db 7 GCGCAACGCGGAGGAGCACTGACCCCGGCGGCGCCAGCCCTCGATATGCAATGAC 66  
 QY 103 TGCCTGCTTGGGAGCAGAGAGTGTGCGGAGGATGAGAGGAGGAGCGGAGCGTCTGCGCA 162  
 Db 67 TGCCTGCTTGGGAGCAGAGAGTGTGCGGAGGATGAGAGGAGGAGCGGAGCGTCTGCGCA 126  
 QY 163 CAGGATCCGGAATCAGATATAGTGGGTTTAACTCCATCAACGCAAAACTCCGCGG 222  
 Db 127 CAGGATCCGGAATCAGATATAGTGGGTTTAACTCCATCAACGCAAAACTCCGCGG 186  
 QY 223 AGCTGGCCCGCTTTTAACTCCGCTCAAGTTTCCCATCCGTAATATGAAGGGGTGG 282  
 Db 187 AGCTGGCCCGCTTTTAACTCCGCTCAAGTTTCCCATCCGTAATATGAAGGGGTGG 246  
 QY 283 ATCTCCGAGCGCTTAACATTCAGAACTCGATGAGGAGGAGGAGGAGGAGGATGAGCC 342  
 Db 247 ATCTCCGAGCGCTTAACATTCAGAACTCGATGAGGAGGAGGAGGAGGAGGATGAGCC 306  
 QY 343 ACCCAACGTCACCTCCCGGCTGAGAGCCCGGCTACACTGATCCAGGGGGTGGCACT 402  
 Db 307 ACCCAACGTCACCTCCCGGCTGAGAGCCCGGCTACACTGATCCAGGGGGTGGCACT 366  
 QY 403 CCGGCGGAGCAGAGCGGGGGTGGGGGCTCTAGAAACCTTACCGGCGGCTTGGCAG 462  
 Db 367 CCGGCGGAGCAGAGCGGGGGTGGGGGCTCTAGAAACCTTACCGGCGGCTTGGCAG 426  
 QY 463 CGCTTAAGCGAGCGCGCGCTCTGCAAGCTTCTGCGCGGAGGATTGCAACCAACGAG 522  
 Db 427 CGCTTAAGCGAGCGCGCGCTCTGCAAGCTTCTGCGCGGAGGATTGCAACCAACGAG 486  
 QY 523 GATGGGACCGGACCGCTTCTGAGGAGGACCACTGTGAGGAGGAGGAGCGGTGCGAG 582  
 Db 487 GATGGGACCGGACCGCTTCTGAGGAGGACCACTGTGAGGAGGAGGAGCGGTGCGAG 546  
 QY 583 ACAAGAGTGTGACTCGGAGTGTGCGGAGGATGAGAGGAGGAGCGGAGCGCTA 642  
 Db 547 ACAAGAGTGTGACTCGGAGTGTGCGGAGGATGAGAGGAGGAGCGGAGCGCTA 606  
 QY 643 ACCGAGCTCCCTCTGCGCGCCGCTCGCAGAGGCGACGTCAGAGGTCCCGGCGCT 702  
 Db 607 ACCGAGCTCCCTCTGCGCGCCGCTCGCAGAGGCGACGTCAGAGGTCCCGGCGCT 666  
 QY 703 CCGTGAACGTT 713  
 Db 667 CCGTGAACGTT 677

RESULT 4  
 BQ187216 655 bp mRNA linear EST 30-APR-2002  
 LOCUS UI-B-BJ1-aj3-e-14-0-UI.r1 UI-B-BJ1 Homo sapiens cDNA clone  
 DEFINITION UI-B-BJ1-aj3-e-14-0-UI 5', mRNA sequence.  
 ACCESSION BQ187216  
 VERSION BQ187216.1 GI:20362767  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Bakayote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homiidae; Homo.  
 1 (bases 1 to 655)

**AUTHORS** Ronaldo, M.F., Lennon, G. and Soares, M.B.  
**TITLE** Normalization and subtraction: two approaches to facilitate gene discovery  
**JOURNAL** Genome Res. 6 (9), 791-806 (1996)  
**PUBMED** 8889548  
**COMMENT** Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Tissue Procurement: Dr. Gregg Hageman  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).  
 The following repetitive elements were found in this cDNA sequence: 296-333, >GC rich#low\_complexity (matched complement)  
 Seq primer: M13 REVERSE

**FEATURES****Source**

Location/Qualifiers  
 1..655  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-B-EJ1-ajz-e-14-0-UI"  
 /tissue\_type="fetal eye, lens, eye anterior segment, optic nerve, retina, Retina Foveal and Macular, RPE and Choroid"  
 /dev\_stage="fetal and adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-B-EJ1"  
 /note="Organ: eye; Vector: pRT73-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-B-EJ1 is a subcloned cDNA library constructed according to Ronaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pRT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (GT)<sub>18</sub> tail. The sequence tags for this library are: fetal eyes, AGAATCAGAA; lens, CGAATAGCGA; eye anterior segment, AATGCCGAT; optic nerve, CCAATAGTG; retina, CCGCG; Retina Foveal and Macular, GTCC; RPE and Choroid, AACTA. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI)."

**ORIGIN**

Query Match 20.8%; Score 649.4; DB 3; Length 655;  
 Best Local Similarity 99.4%; Pred. No. 9,5e-103;  
 Matches 650; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

526 GGGGACCCGACCTTCAGCTTGCAGAGGACCACTGTGAGGCGGTGTCAGAGACA 585  
 Db 2 GGGGACCCGACCTTCAGCTTGCAGAGGACCACTGTGAGGCGGTGTCAGAGACA 61  
 586 CGACGTGTGACTCGGAGTGCCTGTGGAGATGTGACGAGGAGCGGGGACCGCTAAG 645  
 Db 62 CGACGTGTGACTCGGAGTGCCTGTGGAGATGTGACGAGGAGCGGGGACCGCTAAG 121  
 646 GGGGCTCCCTCTGCGCGCCCGTCCGACAGAGGCGACGTGAGGCTCCCGGCGGCTCCG 705  
 Db 122 GGGGCTCCCTCTGCGCGCCCGTCCGACAGAGGCGACGTGAGGCTCCCGGCGGCTCCG 181  
 706 TGGACGTTGGCGGTAGCGCGCGACGATGACGACCATGAGAGCGTTCTGTCGCGCGG 765  
 Db 182 TGGACGTTGGCGGTAGCGCGCGACGATGACGACCATGAGAGCGTTCTGTCGCGCGG 241  
 766 CCCAAGGCGGGGATGCGGGTTAGCAATCTGCGCGCTGAGGGGAGGCTAACGGGCG 825

Db 242 CCCAAGGCGGGATGCGGGTTAGCAATCTGCGCGCTGAGGGGAGGCTAACGGGCG 301  
 826 GGGGCGCGCGGACCCGACGCGGAGCCACCGCGATGCGAGGAGATGCAAGCGCTGC 885  
 Db 302 GGGGCGCGCGGACCCGACGCGGAGCCACCGCGATGCGAGGAGATGCAAGCGCTGC 361  
 886 TGGACGCGCTCAAGAGCGACTGCTGTACCAACCTGCTGAGCCGCTGCTGCT 945  
 Db 362 TGGACGCGCTCAAGAGCGACTGCTGTACCAACCTGCTGAGCCGCTGCTGCT 421  
 946 CGGCGACTCGCAGAACCTGCGGCGAGAGCTGCAAAAGCGCGCAGAGGCGAGAGC 1005  
 Db 422 CGGCGACTCGCAGAACCTGCGGCGAGAGCTGCAAAAGCGCGCAGAGGCGAGAGC 481  
 1006 TGGCGGTGTCCACTGCGCGCGCTGACTGTGTGCTGCGGACCGGCGCTGCGCGG 1065  
 Db 482 TGGCGGTGTCCACTGCGCGCGCTGACTGTGTGCTGCGGACCGGCGCTGCGCGG 541  
 1066 ACGAGCGGCGCGAGTTGAGAGGCGCTGCGGTGCGCTTCTGCGGCTGCGAGCTGCTG 1125  
 Db 542 ACGAGCGGCGCGAGTTGAGAGGCGCTGCGGTGCGCTTCTGCGGCTGCGAGCTGCTG 601  
 1126 AAGCGACATCGACGCTGCTGAGCTGAGGCTGCGCGCGCTTCCGCTGACCGCGC 1179  
 Db 602 AAGCGACATCGACGCTGCTGAGGCTGAGGCTGCGCGCGCTTCCGCTGACCGCGC 655

**RESULT 5**  
 BM671616/c  
 LOCUS  
 DEFINITION UI-B-CQ1-agg-1-04-0-UI .81 UI-B-CQ1 Homo sapiens cDNA clone

**ACCESSION** BM671616  
**VERSION** BM671616.1 GI:18981514  
**KEYWORDS** EST.  
**ORGANISM** Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Hominae; Homo.

**REFERENCE**  
 1 (bases 1 to 685)  
 Ronaldo, M.F., Lennon, G. and Soares, M.B.  
**AUTHORS** Normalization and subtraction: two approaches to facilitate gene discovery  
**TITLE** Genome Res. 6 (9), 791-806 (1996)  
**JOURNAL** 8889548  
**PUBMED**  
**COMMENT** Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Tissue Procurement: Dr. Gregg Hageman  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.resgen.com).  
 Seq primer: M13 Forward

**FEATURES****Source**

Location/Qualifiers  
 1..685  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-B-CQ1-agg-1-04-0-UI"  
 /tissue\_type="optic nerve"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-B-CQ1"  
 /note="Organ: eye; Vector: pRT73-Pac (Pharmacia) with a





Qy	2309	GAAGAGGCTCTCCCAACCAACCGGAAATGGGGTAAAGAAAGCAAGGCTTGGGG	2366
Db	181	GAAGAGGCTCTCCCAACCAACCGGAAATGGGGTAAAGAAAGCAAGGCTTGGGG	240
Qy	2369	TAGGCGCACTGATGTTTAAACAGGCACTTCTCCTCTGAGGCTTATTTTGTGAG	2428
Db	241	TAGGCGCACTGATGTTTAAACAGGCACTTCTCCTCTGAGGCTTATTTTGTGAG	300
Qy	2429	AACTAGACAGATGTTGAACCTCTTTCGAGAGGGCTGGGAATCCTCTTAAAGCAC	2488
Db	301	AACTAGACAGATGTTTGAACCTCTTTCGAGAGGGCTGGGAATCCTCTTAAAGCAC	360
Qy	2489	TTAATCCATTTATCCCGTGGAAATGAGTGTCGTGACGATGAGAGGGCTTGGGAG	2548
Db	361	TTAATCCATTTATCCCGTGGAAATGAGTGTCGTGACGATGAGAGGGCTTGGGAG	420
Qy	2549	CTCCCTGACCCCGCGCTGCCCGCCCTCCGGAGTAATGTGGCATTACTGGCCACAGAG	2608
Db	421	CTCCCTGACCCCGCGCTGCCCGCCCTCCGGAGTAATGTGGCATTACTGGCCACAGAG	480
Qy	2609	GTTTTCAGCCATCACTCTGAGATGGGTAAATGTAAACGCTTTAACTTGGGATTTA	2668
Db	481	GTTTTCAGCCATCACTCTGAGATGGGTAAATGTAAACGCTTTAACTTGGGATTTA	540
Qy	2669	AGAGCTTTTAAAGTATATATCTCTGAAAGAAATAATGACGTAAACACAGCTGTACT	2728
Db	541	AGAGCTTTTAAAGTATATATCTCTGAAAGAAATAATGACGTAAACACAGCTGTACT	600
Qy	2729	ATGAAGCTGTATTTTAAATAAAGACGTGGGCCATGACTCATTA	2774
Db	601	ATGAAGCTGTATTTTAAATAAAGACGTGGGCCATGACTCATTA	646

LOCUS	BO682843	972 bp	mRNA	linear	EST 15-JUL-2002
DEFINITION	AGENCOURT 8495014 NIH_MGC_112 Homo sapiens cDNA clone IMAGE:62951494				
ACCESSION	5', mRNA sequence.				
VERSION	BO682843				
KEYWORDS	BO682843.1	GI:21795522			
SOURCE	EST.				
ORGANISM	Homo sapiens (human)				
	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.				
REFERENCE	1 (bases 1 to 972)				
AUTHORS	NIH-MGC <a href="http://mgc.ncl.nih.gov/">http://mgc.ncl.nih.gov/</a> .				
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)				
JOURNAL	Unpublished (1999)				
COMMENT	Contact: Robert Strausberg, Ph.D. Email: <a href="mailto:cgapbs-remail.nih.gov">cgapbs-remail.nih.gov</a> Tissue Procurement: DCTD/DRP CDNA Library Preparation: Rubin Laboratory CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://image.lnl.gov">http://image.lnl.gov</a> Plate: L1CM2500 row: 0 column: 22 High quality sequence stop: 463.				
FEATURES	location/Qualifiers				
source	1..972				

FEATURES	SOURCE	Location/Qualifiers
1	972	
		/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="taxon:9606"
		/clone="IMAGS:6295149"
		/tissue_type="melanotic melanoma, cell line"
		/lab_host="DH10B (phage-resistant)"
		/clone_1lb="NH MGC_112"
		/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned"

into EcoRI/XhoI sites using the following 5' adaptor: GGCGACGAG(G). Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-CDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies). Note: this is a NIH\_MGC Library."

Query Match	18.5%	Score 577.4	DB 5	Length 972
Beet Local Similarity	96.0%	Prod. No. 3e-90		
Matches 656	Conservative 0	Mismatches 21	Indels 6	Gaps 6
QY 1899	AGTTGGGAGCCACCCCTGACGAGCGTCGCCCTCTTTCGCGGGAAGAGT	1955		
DB 1	AGTTGGGAGCCACCCCTGACGAGCGTCGCCCTCTTTCGCGGGAAGT	60		
QY 1959	GGGTGGGTGAGTGAAGTTTGGAGAGGAGCGCTGTGTTCTATGATGTTGCTGT	2016		
DB 61	GGGTGGGTGAGTGAAGTTTGGAGAGGAGCGCTGTGTTCTATGATGTTGCTGT	120		
QY 2019	TCCCGGACAAAGAAAAATTGCAATCAAAATGTCAGAGCTTTATATCTTATCTTTAG	2078		
DB 121	TCCCGGACAAAGAAAAATTGCAATCAAAATGTCAGAGCTTTATATCTTATCTTTAG	180		
QY 2079	GCTTAATTTTGGAGAGTGTCTGAGACAGTTCTATCAAAAGGCTTCTCTAAGACG	2133		
DB 181	GCTTAATTTTGGAGAGTGTCTGAGACAGTTCTATCAAAAGGCTTCTCTAAGACG	240		
QY 2139	CTACAGCCCTTCTTAGCAGAGTTTATCCATTCTGTCCCAAGACAGCTTAGAAGATTG	2198		
DB 241	CTACAGCCCTTCTTAGCAGAGTTTATCCATTCTGTCCCAAGACAGCTTAGAAGATTG	300		
QY 2139	AGGTCAATACCTTCCCATCTGCGCTCAAGGGGCTGACCCATTTAGAAACAAAGGGGT	2255		
DB 301	AGGTCAATACCTTCCCATCTGCGCTCAAGGGGCTGACCCATTTAGAAACAAAGGGGT	360		
QY 2259	GGTTGAACCTTACTCTACAGGACTTGGATTCAGTGGCACTTGGCTCGGAAAAAGGCT	2318		
DB 351	GGTTGAACCTTACTCTACAGGACTTGGATTCAGTGGCACTTGGCTCGGAAAAAGGCT	420		
QY 2319	CTCCCAAGCACAACCGAGATGGGGGTAAAGAGAAAGACAGAGCTTGGGGTAGGGCA-C	2377		
DB 421	CTCCCAAGCACAACCGAGATGGGGGTAAAGAGAAAGACAGAGCTTGGGGTAGGGCAAC	480		
QY 2378	CTGGGTGTTAAACAGGCACTTCTCTCTCTGCGGGCTTATTTTGTTCAGAACTAAGAC	2433		
DB 481	CTGGGTGTTAAACAGGCACTTCTCTCTCTCTGCGGGCTTATTTTGTTCAGAACTAAGAC	540		
QY 2438	AGAGGTGTTGAA-CCTCCTTGGACAGAGGGGCTGGG-AATCGCTTTAGACA-CTTAATC	2494		
DB 541	AGAGGTGTTGAA-CCTCCTTGGACAGAGGGGCTGGGAAATCCTCTTTAAACACTTAATC	600		
QY 2495	CTATTATAT-CCTCTGGAATGTGGTCTGAGCAGTAAAG-AGGAGCTGCTTTGGACGCTC	2555		
DB 601	CTATTATATCCCTCTGGAATGTGGGCTGGCAAGAAAGGAGGCTGTTTGGAAAGCTCC	660		
QY 2553	CTGACCCCGGCGCTGCGCGCCCTC 2575			
DB 661	CTTGAACCCCGGCGGTGGGCC 683			

RESULT 8					
BM707056					
LOCUS	BM707056	578 bp	mRNA	linear	EST 28-FEB-2002
DEFINITION	UI-B-CR1-adx-b-12-0-UI.r1 UI-B-CR1 Homo sapiens cDNA clone				
	UI-B-CR1-adx-b-12-0-UI 5', mRNA sequence.				
ACCESSION	BM707056				
VERSION	BM707056.1	GI:19020314			
KEYWORDS	EST.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;				



M.Facima Bonaldo. This library was constructed from the same fetus as the fetal lung library, Soares fetal lung NBHL19W."

## ORIGIN

```
Query Match      17.1%; Score 534.8; DB 8; Length 595;
Best Local Similarity 97.7%; Pred. No. 8.2e-83;
Matches 584; Conservative 0; Mismatches 9; Indels 5; Gaps 4;

OY 2129 CTAAGACGGCTACAGCCCTTCTTACGACAGTTTATCCATTGTGCTCCCAAGACGCTAG 2188
    |||||
DB 1 CTAAGACGGCTACAGCCCTTCTTACGACAGTTTATCCATTGTGCTCCCAAGACGCTAG 60

OY 2189 AAGAGATTGAGGTATGATGACCTCCACGCGGCTCAGGGGCTGACCTATTATGAGAAACC 2248
    |||||
DB 61 AAGAGATTGAGGTATGATGACCTCCACGCGGCTCAGGGGCTGACCTATTATGAGAAACC 120

OY 2249 AAGAGGGTGGGTGATGACCTTACTCTCAGGACCTTGAGTCCAGTGGCACACTTGCTGCG 2308
    |||||
DB 121 AAGAGGGTGGGTGATGACCTTACTCTCAGGACCTTGAGTCCAGTGGCACACTTGCTGCG 180

OY 2309 GAAAAGGGCTCTCCCAAGCCACCCGAGATGGGGGTAAAGAGAGACAGAGGCTTGCGG 2368
    |||||
DB 181 GAAAAGGGCTCTCCCAAGCCACCCGAGATGGGGGTAAAGAGAGACAGAGGCTTGCGG 240

OY 2369 TAGGGCCACCTGGGTTTAAACA -GGCACTTCTGCTTCTGCGGGCTTATTTTGTTC 2427
    |||||
DB 241 TAGGGCCACCTGGGTTTAAACAAGCGCACTTCTGCTTCTGCGGGCTTATTTTGTTC 300

OY 2428 GAACTAGACACAGAGTGTGAACTCTCTTTCAGAGAGGGCTGGGAATCTCTTTAGACA 2487
    |||||
DB 301 GAACTAGACACAGAGTGTGAACTCTCTTTCAGAGAGGGCTGGGAATCTCTTTAGACA 360

OY 2488 CTTAATCTTATTTATCCCTGGAAATGTGGTCTGGCCAGTGAAGAGGCTTGCGCA 2547
    |||||
DB 361 CTTAATCTTATTTATCCCTGGAAATGTGGTCTGGCCAGTGAAGAGGCTTGCGCA 420

OY 2548 GCTCCCTGACCCCGGCGCTGCGCCGCTCCGCGGGTAAATGTGGCACTTACGCGCACAGA 2607
    |||||
DB 421 GCTCCCTGACCCCGGCGCTGCGCCGCTCCGCGGGTAAATGTGGCACTTACGCGCACAGA 479

OY 2608 GATTTCAGACCAATCAGCTCTGAGACTGGGTAAATGTGAACAGCTTTAACTTGGGAATTT 2667
    |||||
DB 480 GATTTCAGACCAATCAGCTCTGAGACTGGGTAAATGTGAACAGCTTTAACTTGGGAATTT 539

OY 2668 AAGAACTTTTAAAGTAAATATCTCTGAAAAGAAAATGACGTAAACACAGCGTGT 2725
    |||||
DB 540 AAGAACTTTTAAAGTAAATATCTCTGAAAAGAAAATGACGTAAACACAGCGTGT 594

RESULT 10
BM710194 500 bp mRNA linear EST 28-FEB-2002
LOCUS UI-B-CQ1-agg-1-04-0-UI.r1 UI-B-CQ1 Homo sapiens cDNA clone
DEFINITION UI-B-CQ1-agg-1-04-0-UI 5', mRNA sequence.
ACCESSION BM710184
VERSION BM710194.1 GI:19023452
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 500)
AUTHORS Bonaldo,M.F., Lennon,G. and Soares,M.B.
TITLE Normalization and subtraction: two approaches to facilitate gene
discovery
JOURNAL Genome Res. 6 (9), 791-806 (1996)
PUBMED 8889548
COMMENT Contact: Soares, MB
Coordinated Laboratory for Computational Genomics
University of Iowa
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA
```

Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bentto-soares@iowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
Seq primer: M13 Reverse.  
Location/Qualifiers

FEATURES  
source

```
1..500
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="UI-B-CQ1-agg-1-04-0-UI"
/risue_type="optic nerve"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/clone_lib="UI-B-CQ1"
/note="Organ: eye; Vector: pTT73-Pac (Pharmacia) with a
modified polylinker; Site_1: BcoR I; Site_2: Not I;
UI-B-CQ1 is a normalized cDNA library containing the
following tissue(s): optic nerve. The library was
constructed according to Bonaldo, Lennon and Soares,
Genome Research, 6:791-806, 1996. First strand cDNA
synthesis was primed with an oligo-dT primer containing a
Not I site. Double stranded cDNA was ligated to an EcoR I
adaptor, digested with Not I, and cloned directionally
into pTT73-Pac vector. The oligonucleotide used to prime
the synthesis of first-strand cDNA contains a library tag
sequence that is located between the Not I site and the
(dT)18 tail. The sequence tag for this library is
CCATTAGTG. This library was created for the program, Gene
Discovery in the Visual System, supported by National Eye
Institute (NEI)."
```

## ORIGIN

```
Query Match      16.0%; Score 500; DB 3; Length 500;
Best Local Similarity 100.0%; Pred. No. 9.5e-77;
Matches 500; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2022 CGGACAGAAAATTCGAATCAATATGTCAGAGCTTTATTTATCTTAATCTTTACGGGCC 2081
    |||||
DB 1 CGGACAGAAAATTCGAATCAATATGTCAGAGCTTTATTTATCTTAATCTTTACGGGCC 60

OY 2082 TAAATTTAGAGAGTGTCTGAGAGCACTTCAAAAGGCTTCTCTAAGACGCGCTA 2141
    |||||
DB 61 TAAATTTAGAGAGTGTCTGAGAGCACTTCAAAAGGCTTCTCTAAGACGCGCTA 120

OY 2142 CAGCCCTTCTAGACAGATTATTCATTCTCCCAAGAGCACTAAGAGATTGAGG 2201
    |||||
DB 121 CAGCCCTTCTAGACAGATTATTCATTCTCCCAAGAGCACTAAGAGATTGAGG 180

OY 2202 TCATGACCTCCACATGCGCTCAGGGGCTGACCTATTATGAAAACCAAGAGGGTGGGT 2261
    |||||
DB 181 TCATGACCTCCACATGCGCTCAGGGGCTGACCTATTATGAAAACCAAGAGGGTGGGT 240

OY 2262 TGAACCTACTCTCAGCAGCTTGATTCAGTGCAGCACTTGCGCGAAAAAGGCTCTC 2321
    |||||
DB 241 TGAACCTACTCTCAGCAGCTTGATTCAGTGCAGCACTTGCGCGAAAAAGGCTCTC 300

OY 2322 CCCAGCCACCCGAGATGGGGGTAAAGAGAGACAGAGGCTTGGGGTAAAGGCTCTG 2381
    |||||
DB 301 CCCAGCCACCCGAGATGGGGGTAAAGAGAGACAGAGGCTTGGGGTAAAGGCTCTG 360

OY 2382 TGTTTAAACAGGCACTTCTCTGCGGGCTTATTTTGTTCAGAACTAGACAGAG 2441
    |||||
DB 361 TGTTTAAACAGGCACTTCTCTCTGCGGGCTTATTTTGTTCAGAACTAGACAGAG 420

OY 2442 TGTTTGAACCTCTTTGACAGAGGGCTGGGAATCTCTTTAGACACTTAATCTTATTA 2501
    |||||
DB 421 TGTTTGAACCTCTTTGACAGAGGGCTGGGAATCTCTTTAGACACTTAATCTTATTA 480
```

QY 2502 TCCCTGGAATGTCGTCT 2521  
 Db 481 TCCCTGGAATGTCGTCT 500

RESULT 11  
 AM302149 476 bp mRNA linear EST 18-JAN-2000  
 LOCUS AM302149/c  
 DEFINITION x001f06.x1 NCI CGAP Kid11 Homo sapiens CDNA clone IMAGE:2768387 3' similar to confine TAR1 repetitive element ;, mRNA sequence.

ACCESSION AM302149  
 VERSION AM302149.1 GI:6711826  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
 TITLE 1 (bases 1 to 476)  
 JOURNAL NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 COMMENT National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgaps-remail.nih.gov  
 Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Bamber-Buck, M.D., Ph.D.  
 CDNA Library Preparation: M. Bento Soares, Ph.D.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/INMUT at: www-bio.lnln.gov/bdrrp/image/image.html

FEATURES  
 source  
 Possible reversed clone: polyT not found  
 Seq primer: -40UP from Gibco  
 High quality sequence stop: 457.  
 Location/Qualifiers  
 1..476  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2768387"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP Kid11"  
 /note="Organ: kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; Plasmid DNA from the normalized library NCI CGAP Kid3 was prepared, and as circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneIDs 1322376-1323911, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN  
 Query Match 15.1%; Score 472.8; DB 1; Length 476;  
 Best Local Similarity 99.6%; Pred. No. 5.1e-72;  
 Matches 474; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 716 CGGTAGGCGCGAGGAGTCAAGCAATGAAAGAGCGTTCGTCGCGCGCGCGCAAGGCGG 775  
 Db 476 CGGTAGGCGCGAGGAGTCAAGCAATGAAAGAGCGTTCGTCGCGCGCGCGCAAGGCGG 417

QY 776 GGATGGGGGTTAGCACATCTCTGCGCGCTGAGGGGGAGGCTAAACGGAGCGCGCGCGCG 835  
 Db 416 GGATGGGGGTTAGCACATCTCTGCGCGCTGAGGGGGAGGCTAAACGGAGCGCGCGCGCG 357

QY 836 GGGCCAGCGCGAGCGCGAGTGGGAGGAGAGAGTGCACAGCGCGCTCTGAGCGGGCT 895  
 Db 356 GGGCCAGCGCGAGCGCGAGTGGGAGGAGAGAGTGCACAGCGCGCTCTGAGCGGGCT 297

QY 896 CAAAGAGAGACTGCGTGTCTACCAACCACTGTGTGACCGTGTGAGTGTGCGGAGCTC 955  
 Db 296 CAAAGAGAGACTGCGTGTCTACCAACCACTGTGTGACCGTGTGAGTGTGCGGAGCTC 237

QY 956 GCGAAGCTGCGCGGAGAGTGTGCAAAAGACGCGCGCAAGAGCGGAGAGTGTGTC 1015  
 Db 236 GCGAAGCTGCGCGGAGAGTGTGCAAAAGACGCGCGCAAGAGCGGAGAGTGTGTC 1177

QY 1016 CACCTGCGCGCGGAGTGTGCTGTGTGTCGCGGAGCGGCGGCGCTGAGCGAGCGCGC 1075  
 Db 176 CACCTGCGCGCGGAGTGTGCTGTGTGTCGCGGAGCGGCGGCGCTGAGCGAGCGCGC 117

QY 1076 CGAGTTGAGCGGCTGTGCGGTGCGCTTCTGCGGCTGCGTGTGAGCGGAGCAT 1135  
 Db 116 CGAGTTGAGCGGCTGTGCGGTGCGCTTCTGCGGCTGCGTGTGAGCGGAGCAT 57

QY 1136 GCGAGCGCTGCGTGTGAGAGTGTGCGCGCGCGCTTCCGCTGTGACGCGCGCGGAGCGCG 1191  
 Db 56 GCGAGCGCTGCGTGTGAGAGTGTGCGCGCGCGCTTCCGCTGTGACGCGCGCGGAGCGCG 1

RESULT 12  
 BM673230/c 480 bp mRNA linear EST 27-FEB-2002  
 LOCUS BM673230/c  
 DEFINITION UI-E-CRI-idx-b-12-0-UI.s1 UI-E-CRI Homo sapiens CDNA clone  
 UI-E-CRI-idx-b-12-0-UI 3', mRNA sequence.

ACCESSION BM673230  
 VERSION BM673230.1 GI:18983128  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE  
 AUTHORS Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
 TITLE 1 (bases 1 to 480)  
 JOURNAL Normalization and subtraction: two approaches to facilitate gene discovery  
 PubMed 8889548  
 Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Tissue Procurement: Dr. Gregg Hageman  
 CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research Genetics (www.regen.com).  
 Seq primer: M13 Forward  
 POLVA=Yes.

FEATURES  
 source  
 Location/Qualifiers  
 1..480  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-E-CRI-idx-b-12-0-UI"  
 /clone="UI-E-CRI-idx-b-12-0-UI"  
 /tissue\_type="eye anterior segment"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_lib="UI-E-CRI"  
 /note="Organ: eye; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: EcoR I; Site 2: Not I; UI-E-CRI is a normalized CDNA library containing the following tissue(s): eye anterior segment. The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. First strand cDNA synthesis was primed with an oligo-dT primer containing a

Not 1 site. Double stranded cDNA was ligated to an EcoR I adaptor, digested with Not I, and cloned directionally into pT73-Pac vector. The oligonucleotide used to prime the synthesis of first-strand cDNA contains a library tag sequence that is located between the Not I site and the (dT)18 tail. The sequence tag for this library is AATGCCGCAAT. This library was created for the program, Gene Discovery in the Visual System, supported by National Eye Institute (NEI).  
TAG\_TISSUE=human eye anterior segment  
TAG\_LIB=UI-B-CRI  
TAG\_SEQ=AATGCCGCAAT

## ORIGIN

Query Match 14.8%; Score 461.8; DB 3; Length 480;  
Best Local Similarity 99.6%; Pred. No. 4.2e-70;  
Matches 463; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2310 AAAAGGGCTCTCCAGCCAGCCGAGATGGGGTAAAGAAAGAGAGAGGCTTGGAGT 2369  
DB AAAAGGGCTCTCCAGCCAGCCGAGATGGGGTAAAGAAAGAGAGAGGCTTGGAGT 421

QY 2370 AGGGCCACCTGGTGTAAACAGGCACTTCTCTCTCTGGGGCTTAATTTTGTTCAGA 2429  
DB AGGGCCACCTGGTGTAAACAGGCACTTCTCTCTCTGGGGCTTAATTTTGTTCAGA 361

QY 2430 ACTAGACCAAGATGTTTGAACCTCTTTTGACAGAGGGCTGGAAATCCTTTAGGCACT 2489  
DB ACTAGACCAAGATGTTTGAACCTCTTTTGACAGAGGGCTGGAAATCCTTTAGGCACT 301

QY 2490 TAATCTTAATTTATCCCTGGAATGTCGTGCGCAGTAGAGAGGGCTGGCTTTGGCAGC 2549  
DB TAATCTTAATTTATCCCTGGAATGTCGTGCGCAGTAGAGAGGGCTGGCTTTGGCAGC 241

QY 2550 TCCCTGACCCCGCGCTGCGCCCTCCCGGGTAAATGTGCACTTAATCTGCGCCACAGAG 2609  
DB TCCCTGACCCCGCGCTGCGCCCTCCCGGGTAAATGTGCACTTAATCTGCGCCACAGAG 181

QY 2610 TTTTGAAGCAATCAGCTCTGAGACTGGGTAGAAATGTAAACAGCTTAACTTGGAGTTTAA 2669  
DB TTTTGAAGCAATCAGCTCTGAGACTGGGTAGAAATGTAAACAGCTTAACTTGGAGTTTAA 121

QY 2670 GAAGCTTTAAAGGTAATATCTCTGAAGAAATAATGACGTAAACAAGGCTTACATA 2729  
DB GAAGCTTTAAAGGTAATATCTCTGAAGAAATAATGACGTAAACAAGGCTTACATA 61

QY 2730 TGAAGCTGTATTATTAAATAAGAACGCTGGGCCATGAATCATTA 2774  
DB TGAAGCTGTATTATTAAATAAGAACGCTGGGCCATGAATCATTA 16

RESULT 13  
BMS62937 1376 bp mRNA linear EST 20-FEB-2002  
LOCUS BMS62937  
DEFINITION AGENCOURT\_6566935 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:5736787  
5' mRNA sequence.  
ACCESSION BMS62937  
VERSION BMS62937.1 GI:18809436  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS 1 (bases 1 to 1376)  
TITLE NIH-MGC http://mgi.nci.nih.gov/.  
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
COMMENT Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)

DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov  
Plate: LLM12745 row: n column: 20  
High quality sequence stop: 425.

## FEATURES

source

1..1376  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5736787"  
/tissue\_type="duodenal adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_88"  
/note="Organ: small intestine; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.767 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 14.3%; Score 445.2; DB 3; Length 1376;  
Best Local Similarity 96.9%; Pred. No. 2.7e-67;  
Matches 475; Conservative 0; Mismatches 13; Indels 2; Gaps 2;

QY 453 CCTTTGGAGGCTTAAGGCGGAGCGCGGCTCTGACGCTTGCTTGCCTCCGAGTTGGC 512  
DB CCTTTGGAGGCTTAAGGCGGAGCGCGGCTCTGACGCTTGCTTGCCTCCGAGTTGGC 60

QY 513 ACCACGAGAGATGGGGACGCGCACCTCAGCTTGCAGAGAGGACCAACGCTGAGGCCAGG 572  
DB ACCACGAGAGATGGGGACGCGCACCTCAGCTTGCAGAGAGGACCAACGCTGAGGCCAGG 120

QY 573 CGGTGACAGACACAGACGTGTGACTCGAGAGTGCCTTGGAGAGATGACAGAGAGCGG 632  
DB CGGTGACAGACACAGACGTGTGACTCGAGAGTGCCTTGGAGAGATGACAGAGAGCGG 180

QY 633 GGGACCGCTAACCGGGGCTCCTCTTGCCTGCGCCCTCCGACAGAGCGCACCTGAGAGTCC 692  
DB GGGACCGCTAACCGGGGCTCCTCTTGCCTGCGCCCTCCGACAGAGCGCACCTGAGAGTCC 240

QY 693 CGGGCGGGCTCCGTGAGAGTTGGCGGTAGCGCCGAGCAGTCAAGACATGAAGAGCGT 752  
DB CGGGCGGGCTCCGTGAGAGTTGGCGGTAGCGCCGAGCAGTCAAGACATGAAGAGCGT 300

QY 753 TCGTGCCTGCGCGCCCAAGGCGGGATGGGGTTAGCCATCTCTGCGCGCTGAAGGGG 812  
DB TCGTGCCTGCGCGCCCAAGGCGGGATGGGGTTAGCCATCTCTGCGCGCGCTGAAGGGG 360

QY 813 AGGCTAACGGGCGCGGGCGCGCGGCCCAAGCCGACCGGATGCGAGAGGAGAG 872  
DB AGGCTAACGGGCGCGGGCGCGCGGCCCAAGCCGACCGGATGCGAGAGGAGAG 420

QY 873 TGCAA-GGCGCTGTGAGCGGGCTCAACAAGACAGTGGTGTACCAACCACTGG-TGC 930  
DB TGCAA-GGCGCTGTGAGCGGGCTTCAACAAGACAGTGGTGTACCAACCACTGGCGGC 480

QY 931 TGACCGTGG 940  
DB TGACCGTGG 490

RESULT 14  
A1375213/c 455 bp mRNA linear EST 18-MAR-1999  
LOCUS A1375213  
DEFINITION tc10f06.x1 Soares\_NbHMPu\_51 Homo sapiens cDNA clone IMAGE:2063459  
3' mRNA sequence.  
ACCESSION A1375213  
VERSION A1375213.1 GI:4175203  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens





```
Qy 1568 GACACCCGACGCGCCCTGCTGCTGCGCCCTCCCTCCCTGGA----- 1608
Db 492 TGCAACCGACGCGCCCGGAGCTACCGCTCCCTCCCTGAGGCAACCCCTCCAGAACAA 433
Qy 1609 -----GAAAGACTGGGATGGGTGTGGGGGTCTGGCTGTGCAAGGGAGTGTCTT 1660
Db 432 CTGCACTGCTAGGACTAGGGGTGGAGTGGGATGTGGCATGTTCAGGAGAGCTGTCT 373
Qy 1661 AAAACCCGCTGTGTCATGGGTACACGGCCGTTTCCAGTGCACATCTGCTGGGACAGAC 1720
Db 372 CGTGCCT--TATACGTGTAGGTATAGCACGTTCAAGCACACATTTGCTAGGACAGAAC 316
Qy 1721 ACGGTTTCTCTTGTGCTGAGCCCGGAGAAATTAACTTTGCGCCGCGCGTCAAGGCATTAC 1780
Db 315 ATGG-CTTCTCTTGTGCTGAGCCGAGGAAAGTTAATTTGCGCGGCGCAGAGGCATTAT 257
Qy 1781 CGCTAACGTCTGACAGAGCTTTATTCCTATTATATGAAAAACGT-CAAGTGAACCTTAG 1839
Db 256 CGCTAATGTATGACAGAGCTTTATCTCTATTATATGAAAAACGTCCACAGTGAACCGAG 197
Qy 1840 ATCCCTCCGAGTTATGAGTTAACACATGTGCTGTGGGGCGTCTTTACAGGAGTCCGA 1899
Db 196 ATTCCTCCAGTTATATGAGTTAACCGCTGTGTGCTGGGGCGTGTTTACACTGAATCTGG 137
Qy 1900 GTTCGCTGCCACCCCTGCGCAGCGTCCGCCCTT-TCGCGTGGGACAGTTT-----GAA 1953
Db 136 GCTTGTCTCTCTCGCTCCCTGTCAAGTATCCCAATCTCTGAGTGGGGCAGTTTGGGGGAA 77
Qy 1954 AAGTGGGTGGGGTGGAGTGAAGTTTGGAGAGGAGCGCTGTTGGTCTTATGTGGTTG 2011
Db 76 AAAAAGGTGAGGTGAATGAAGTGTGAGAAAGTGTCTATTTGTGTGTGTGGGTG 19
```

Search completed: May 11, 2006, 05:27:41  
Job time : 11794 secs

**THIS PAGE BLANK (uspto)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: May 11, 2006, 01:38:14 / Search time 530 Seconds  
(without alignments)  
10470.844 Million cell updates/sec

Title: US-10-760-320A-102

Perfect score: 3122

Sequence: 1 actagaggtggttagcgccg.....acagagcaagactctctc 3122

Scoring table:

IDENTITY\_NUC

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database:

Issued Patents NA:  
1: /cgn2\_6/prodata/1/ina/1 COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5 COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/H COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/PP COMB.seq:\*  
7: /cgn2\_6/prodata/1/ina/RE COMB.seq:\*  
8: /cgn2\_6/prodata/1/ina/RE COMB.seq:\*  
9: /cgn2\_6/prodata/1/ina/BACKFILES1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	181.4	5.8	601	US-09-949-016-196942	Sequence 196942, A
2	181.4	5.8	109250	US-09-949-016-12530	Sequence 12530, A
3	181.4	5.8	109251	US-09-949-016-17321	Sequence 17321, A
4	177.2	5.7	26760	US-09-949-016-15894	Sequence 15894, A
5	176.4	5.7	601	US-09-949-016-40113	Sequence 40113, A
6	176.4	5.7	601	US-09-949-016-138968	Sequence 138968, A
7	176.4	5.7	30371	US-09-949-016-15395	Sequence 15395, A
8	176.4	5.7	41589	US-09-949-016-12867	Sequence 12867, A
9	176.4	5.7	41593	US-09-949-016-15666	Sequence 15666, A
10	176	5.6	601	US-09-949-016-40112	Sequence 40112, A
11	176	5.6	601	US-09-949-016-138967	Sequence 138967, A
12	175.6	5.6	93778	US-09-949-016-15096	Sequence 15096, A
13	174.8	5.6	23766	US-09-949-016-13569	Sequence 13569, A
14	174.8	5.6	74545	US-09-949-002-606	Sequence 606, App
15	174.8	5.6	125192	US-09-949-016-14120	Sequence 14120, A
16	174.8	5.6	131860	US-09-949-002-730	Sequence 730, App
17	174.2	5.6	98828	US-09-949-016-16630	Sequence 16630, A
18	174	5.6	5371	US-09-949-016-16996	Sequence 16996, A
19	174	5.6	152481	US-09-949-016-12521	Sequence 12521, A
20	174	5.6	152798	US-09-949-016-12775	Sequence 12775, A
21	174	5.6	152822	US-09-949-016-17518	Sequence 17518, A
22	174	5.6	152822	US-09-949-016-17519	Sequence 17519, A
23	173.8	5.6	18591	US-09-949-016-14719	Sequence 14719, A
24	173.2	5.5	601	US-09-949-016-81759	Sequence 81759, A

C	25	173.2	5.5	48018	3	US-09-949-016-17573	Sequence 17573, A
	26	173.2	5.5	67386	3	US-09-949-016-16519	Sequence 16519, A
	27	173.2	5.5	146095	3	US-09-949-016-12872	Sequence 12872, A
	28	173.2	5.5	146104	3	US-09-949-016-13239	Sequence 13239, A
	29	173	5.5	59828	3	US-09-949-016-16238	Sequence 16238, A
	30	172.8	5.5	36816	3	US-09-949-016-16218	Sequence 16218, A
	31	172.4	5.5	35688	3	US-09-949-016-16873	Sequence 16873, A
	32	172.4	5.5	49848	3	US-09-949-016-15675	Sequence 15675, A
	33	172.4	5.5	74096	3	US-09-949-016-11785	Sequence 11785, A
	34	172.4	5.5	74097	3	US-09-949-016-15239	Sequence 15239, A
	35	172.2	5.5	18448	3	US-09-949-016-15345	Sequence 15345, A
	36	172	5.5	60376	3	US-09-949-016-12423	Sequence 12423, A
	37	171.8	5.5	282	2	US-08-133-629-8	Sequence 8, Appl1
	38	171.8	5.5	601	3	US-09-949-016-12655	Sequence 12655, A
	39	171.8	5.5	35337	3	US-09-949-016-117249	Sequence 117249, A
	40	171.6	5.5	601	3	US-09-949-016-42392	Sequence 42392, A
	41	171.6	5.5	38368	3	US-09-949-016-12958	Sequence 12958, A
	42	171.6	5.5	58782	3	US-09-949-016-16851	Sequence 16851, A
	43	171.6	5.5	107941	3	US-09-949-016-14206	Sequence 14206, A
	44	171.6	5.5	109378	3	US-09-949-016-12391	Sequence 12391, A
	45	171.4	5.5	20347	3	US-09-949-016-16752	Sequence 16752, A

## ALIGNMENTS

```
RESULT 1
US-09-949-016-196942/C
; Sequence 196942, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 196942
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-196942

Query Match      5.8%; Score 181.4; DB 3; Length 601;
Best Local Similarity 88.5%; Pred. No. 2.7e+29;
Matches 208; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
```

2888 TGAGCAGGTGATACCTTGAGCCAGAGTTCGACCAAGCTGCGCAATAGAGAA 2947  
|||||  
536 TGAGCAGGTGATACCTTGAGCTCGAGATTCGAGCCAGCTGCGCAATAGAGAA 477  
|||||  
2948 CCCGATCTCTAATAAATAATTAATTTGGCCGGGCAATGCTGCTGCTGCTC 3007  
|||||  
476 CCCGATCTCTAATAAATAATTAATTTAGTGGGTGTGTGGCAATGCTGTAATTC 417  
|||||  
3008 CAGCTACTCTGAGAGGTTGAGCGAGAGTCTGTAATGACGAGGTGAGATTGCAT 3067  
|||||  
416 CAGCTACTCTGAGAGGTTGAGCGAGAGTCTGTAATGACGAGGTGAGATTGCAT 357  
|||||  
3068 GAG-CAAGATTGTCACATGCACTCAGCTGCGGCAAGAGCAAGACTGTCT 3121  
|||||  
Db GAGCAAGATTGTCACATGCACTCAGCTGCGGCAAGAGCAAGACTGTCT 302  
|||||

RESULT 2

```
US-09-949-016-12530
; Sequence 12530, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12530
; LENGTH: 109250
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(109250)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12530
```

```
Query Match          5.8%; Score 181.4; DB 3; Length 109250;
Best Local Similarity 88.5%; Pred. No. 1.5e-28;
Matches 208; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
```

```
QY 2888 TGAGCAGGTGATCACCCTGAGGCCAGAGTTCCAGACCAAGCTGGCCCACTAGCGAAA 2947
DB 78099 TGAGCAGGTGATCACCCTGAGGCCAGAGTTCCAGACCAAGCTGGCCCACTAGCGAAA 78158
QY 2948 CCCGATCTCTACTAAAAATATATATATATATATATATATATATATATATATATATAT 3007
DB 78159 CCCGATCTCTACTAAAAATATATATATATATATATATATATATATATATATATATAT 78218
QY 3008 CAGCTACTCGGAGGTTGAGGAGGAGGAGTCCCTTGAATGAGAGAGTGGATTCAT 3067
DB 78219 CAGCTACTCGGAGGTTGAGGAGGAGGAGTCCCTTGAATGAGAGAGTGGATTCAT 78278
QY 3068 GAG-CAAGATTGTGCGACTGCTGAGCTGGGCAACAGACCAAGCTCTGTCT 3121
DB 78279 GAGCAAGATTGTGCGACTGCTGAGCTGGGCAACAGACCAAGCTCTGTCT 78333
```

## RESULT 3

```
US-09-949-016-17321
; Sequence 17321, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17321
; LENGTH: 109251
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
```

```
; NAME/KEY: misc.feature
; LOCATION: (1)...(109251)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17321
```

```
Query Match          5.8%; Score 181.4; DB 3; Length 109251;
Best Local Similarity 88.5%; Pred. No. 1.5e-28;
Matches 208; Conservative 0; Mismatches 26; Indels 1; Gaps 1;
```

```
QY 2888 TGAGCAGGTGATCACCCTGAGGCCAGAGTTCCAGACCAAGCTGGCCCACTAGCGAAA 2947
DB 78099 TGAGCAGGTGATCACCCTGAGGCCAGAGTTCCAGACCAAGCTGGCCCACTAGCGAAA 78158
QY 2948 CCCGATCTCTACTAAAAATATATATATATATATATATATATATATATATATATATAT 3007
DB 78159 CCCGATCTCTACTAAAAATATATATATATATATATATATATATATATATATATATAT 78218
QY 3008 CAGCTACTCGGAGGTTGAGGAGGAGGAGTCCCTTGAATGAGAGAGTGGATTCAT 3067
DB 78219 CAGCTACTCGGAGGTTGAGGAGGAGGAGTCCCTTGAATGAGAGAGTGGATTCAT 78278
QY 3068 GAG-CAAGATTGTGCGACTGCTGAGCTGGGCAACAGACCAAGCTCTGTCT 3121
DB 78279 GAGCAAGATTGTGCGACTGCTGAGCTGGGCAACAGACCAAGCTCTGTCT 78333
```

## RESULT 4

```
US-09-949-016-15894
; Sequence 15894, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15894
; LENGTH: 26760
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15894
```

```
Query Match          5.7%; Score 177.2; DB 3; Length 26760;
Best Local Similarity 87.6%; Pred. No. 7.4e-28;
Matches 205; Conservative 0; Mismatches 28; Indels 1; Gaps 1;
```

```
QY 2889 GAGGCAAGTGTGATCACCCTGAGGCCAGAGTTCCAGACCAAGCTGGCCCACTAGCGAAA 2948
DB 21533 GAGGCAAGTGTGATCACCCTGAGGCCAGAGTTCCAGACCAAGCTGGCCCACTAGCGAAA 21592
QY 2949 CCGATCTCTACTAAAAATATATATATATATATATATATATATATATATATATATAT 3008
DB 21593 CCGATCTCTACTAAAAATATATATATATATATATATATATATATATATATATATAT 21652
QY 3009 AGCTACTCGGAGGTTGAGGAGGAGGAGTCCCTTGAATGAGAGAGTGGATTCAT 3068
DB 21653 AGCTACTCGGAGGTTGAGGAGGAGGAGTCCCTTGAATGAGAGAGTGGATTCAT 21712
QY 3069 AG-CAAGATTGTGCGACTGCTGAGCTGGGCAACAGACCAAGCTCTGTCT 3121
DB 21713 AGCGAAGATTGTGCGACTGCTGAGCTGGGCAACAGACCAAGCTCTGTCT 21766
```

## RESULT 5





Db 27160 CAAACTCTGTCTC 27173

## RESULT 8

```
US-09-949-016-12867/c
; Sequence 12867, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12867
; LENGTH: 41589
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(41589)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12867
```

Query Match 5.7%; Score 176.4; DB 3; Length 41589;  
Best Local Similarity 83.5%; Pred. No. 1.3e-27;  
Matches 212; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

```
Qy 2870 CCCCACTTTTCTTTTGTAGGAGATGATCACTGAGGCGAGAGTTGAGACCAAC 2929
    |||||
Db 8267 CCCAACATTTGGGAATCCGAGCGGTGATCATCTTGAGGCCAGAGTTCAAGACCAAC 8208
    |||||
Qy 2930 CTGGCCAAATGCGAAACCCGATCTCTAATAAATAAATAATTTGCGCGGCAATG 2989
    |||||
Db 8207 CTGGCCAAATGCGAAATTCCTCTAATAAATAAATAATTTGCGCGGCAATG 8148
    |||||
Qy 2990 GGGGATGCTGTGTGCTCCAGCTACTCGGAGGTTGAGGAGAGAGTGGCTTGATGCA 3049
    |||||
Db 8147 GGGGACACCTGTAAATCCAGCTACTCGGAGGCTGAGGAGAAACCGCTTGAACCA 8088
    |||||
Qy 3050 GAGGTGAGAGTTGAGATGAG-CAAGATTGTGCCACTGCACTCGGCGAACAAG 3108
    |||||
Db 8087 GAGGCGAGAGTTGAGATGAGCAAGCAATCAAGCACTGCACTCGGCGAACAAG 8028
    |||||
Qy 3109 CAAGACTCTGTCTC 3122
    |||||
Db 8027 CAAGACTCTGTCTC 8014
    |||||
```

```
RESULT 9
US-09-949-016-15666/c
; Sequence 15666, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
```

```
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15666
; LENGTH: 41593
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(41593)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15666
```

Query Match 5.7%; Score 176.4; DB 3; Length 41593;  
Best Local Similarity 83.5%; Pred. No. 1.3e-27;  
Matches 212; Conservative 0; Mismatches 41; Indels 1; Gaps 1;

```
Qy 2870 CCCCACTTTTCTTTTGTAGGAGATGATCACTGAGGCGAGAGTTGAGACCAAC 2929
    |||||
Db 8267 CCCAACATTTGGGAATCCGAGCGGTGATCATCTTGAGGCCAGAGTTCAAGACCAAC 8208
    |||||
Qy 2930 CTGGCCAAATGCGAAACCCGATCTCTAATAAATAAATAATTTGCGCGGCAATG 2989
    |||||
Db 8207 CTGGCCAAATGCGAAATTCCTCTAATAAATAAATAATTTGCGCGGCAATG 8148
    |||||
Qy 2990 GGGGATGCTGTGTGCTCCAGCTACTCGGAGGTTGAGGAGAGAGTGGCTTGATGCA 3049
    |||||
Db 8147 GGGGACACCTGTAAATCCAGCTACTCGGAGGCTGAGGAGAAACCGCTTGAACCA 8088
    |||||
Qy 3050 GAGGTGAGAGTTGAGATGAG-CAAGATTGTGCCACTGCACTCGGCGAACAAG 3108
    |||||
Db 8087 GAGGCGAGAGTTGAGATGAGCAAGCAATCAAGCACTGCACTCGGCGAACAAG 8028
    |||||
Qy 3109 CAAGACTCTGTCTC 3122
    |||||
Db 8027 CAAGACTCTGTCTC 8014
    |||||
```

```
RESULT 10
US-09-949-016-40112/c
; Sequence 40112, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 40112
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-40112
```

Query Match 5.6%; Score 176; DB 3; Length 601;  
Best Local Similarity 83.1%; Pred. No. 3.8e-28;  
Matches 211; Conservative 1; Mismatches 41; Indels 1; Gaps 1;

```
Qy 2870 CCCCACTTTTCTTTTGTAGGAGATGATCACTGAGGCGAGAGTTGAGACCAAC 2929
    |||||
Db 512 CCCAACATTTGGGAATCCGAGCGGTGATCATCTTGAGGCCAGAGTTCAAGACCAAC 453
    |||||
Qy 2930 CTGGCCAAATGCGAAACCCGATCTCTAATAAATAAATAATTTGCGCGGCAATG 2989
    |||||
```



```

Db      4585 CCCACATTTTGGGGGGCCGAGGTGGGGGATCACCTGAGTCAGAGTTGTAACCAACG 4526
Qy      2930 CTGGCCCAATAGCGAAACCCGATCTCTAATAAAATATAAAATTTGGCCGGCATGGT 2989
Db      4525 CTGGTCAATGTGGCAAAACACCATCTCTAATAAAATATAAAATTTCCGGGTATGGT 4466
Qy      2990 GGGCGATGCTGTGTGCTCCAGTACTCGGGAGGTGAGGCAAGAGTCGCTTGATGCA 3049
Db      4465 GGCATGTGCTGTAGTCCAGCTACTTGGAGGCTGAGGCAAGAGTCACTTGAACCTG 4406
Qy      3050 GAGGTGAGGTGCAATGAG-CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAG 3108
Db      4405 GAGGTGAGGTGCAATGAGTCAAGATTGTGCCACTTGCATCTCCAGCCTGGGCAACAGAG 4346
Qy      3109 CAAGACTCTGTCTC 3122
Db      4345 CAAGACTCTGTCTC 4332

```

## RESULT 14

```

US-09-949-002-606/c
; Sequence 606, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 606
; LENGTH: 74545
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-606

```

```

Query Match      5.6%; Score 174.8; DB 3; Length 74545;
Best Local Similarity 83.1%; Pred. No. 3.3e-27;
Matches 211; Conservative 0; Mismatches 42; Indels 1; Gaps 1;

Qy      2870 CCCACTTTTCTTTTCTTTTGAAGGAGGTGATCACTGAGGCCAGAGTTGAGACCAAGC 2929
Db      11479 CCCACAGTTTAAAGGCTGAAGTGGAGATCACTGAGGTCAAGAGTTCAAGACCAAGC 11420
Qy      2930 CTGGCCCAATAGCGAAACCCGATCTCTAATAAAATATAAAATTTGGCCGGCATGGT 2989
Db      11419 CTGGCCCAATAGAGAAACCCGATCTCTAATAAAATATAAAATTTCCAGAGTGGT 11360
Qy      2990 GGGCGATGCTGTGTGCTCCAGTACTCGGGAGGTGAGGCAAGAGTCGCTTGATGCA 3049
Db      11359 GACAGGCGCTGTATCCAGTACTCGGAGGCTGAGGCAAGAGTCTTGAACCGG 11300
Qy      3050 GAGGTGAGGTGCAATGAG-CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAG 3108
Db      11299 GAGGCGGAGGTGCAATGAGTCAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAG 11240
Qy      3109 CAAGACTCTGTCTC 3122
Db      11239 CAAGACTCTGTCTC 11226

```

## RESULT 15

```

US-09-949-016-14120/c
; Sequence 14120, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

```

```

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14120
; LENGTH: 125192
; TYPE: DNA
; ORGANISM: Human
; NAME/KEY: misc feature
; LOCATION: (1)-(125192)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14120

```

```

Query Match      5.6%; Score 174.8; DB 3; Length 125192;
Best Local Similarity 83.1%; Pred. No. 4e-27;
Matches 211; Conservative 0; Mismatches 42; Indels 1; Gaps 1;

Qy      2870 CCCACTTTTCTTTTCTTTTGAAGGAGGTGATCACTGAGGCCAGAGTTGAGACCAAGC 2929
Db      67067 CCCAGCACTTTTGAAGGCTGAGGCAAGGAGTCACTGAGGTGAGAGTTCAAGACCAAG 67008
Qy      2930 CTGGCCCAATAGCGAAACCCGATCTCTAATAAAATATAAAATTTGGCCGGCATGGT 2989
Db      67007 CTGGCCCAATAGTGAAGACCTGTCTCTAATAAAATATAAAATTTAGCCGGCGTGGT 66948
Qy      2990 GGGCGATGCTGTGTGCTCCAGTACTCGGGAGGTGAGGCAAGAGTCGCTTGATGCA 3049
Db      66947 GGTGATCTCTATATCCCACTACTCGGAAGGTTGAGGCAAGAGTCACTTGAACCCA 66888
Qy      3050 GAGGTGAGGTGCAATGAG-CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAG 3108
Db      66887 GAGGCGGAGGTGCAATGAGTCAAGATTGTGCCACTGCACTCCAGGCAACAGAG 66828
Qy      3109 CAAGACTCTGTCTC 3122
Db      66827 TAAAGCTCTGTCTC 66814

```

Search completed: May 11, 2006, 06:02:58  
Job time : 532 sec



Db	Sequence	Score	DB	Length
Db	181 CAAAGAGAGACTGGGTGTACCAACACTGGGTGTGAACCTGGGTGTGCTGGCGGACATC	240		
Qy	956 GCAGAACTGGCGGCGAGAGCTGCAGAAAAGACGGCGCAGAAAGCGCAGAGCTGGCGGTGTC	1015		
Db	241 GCAGAACTGGCGGCGAGAGCTGCAGAAAAGACGGCGCAGAAAGCGCAGAGCTGGCGGTGTC	300		
Qy	1016 CACCTGGCGCGCGGCTGAGCTGTGTGTGGCGCAGCGGGGCTGTGGCGCGCGAGCGCGC	1075		
Db	301 CACCTGGCGCGCGGCTGAGCTGTGTGTGGCGCAGCGGGGCTGTGGCGCGCGAGCGCGC	360		
Qy	1076 CGAGTTCAGACGGCTCTGGGTGTGACCTTCTTCGGGCTGTGACCTGTGTGAAAGCGACAT	1135		
Db	361 CGAGTTCAGACGGCTCTGGGTGTGACCTTCTTCGGGCTGTGACCTGTGTGAAAGCGACAT	420		
Qy	1136 GCGAGGCTCGCTGTGAGCTGTGGGCGCGCGTTCCTCGGCTGTGACAGCGCGCGCGGACACCGCTGGT	1195		
Db	421 GCGAGGCGCGCTGTGAGCTGTGGGCGCGCGTTCCTCGGCTGTGACAGCGCGCGCGGACACCGCTGGT	480		
Qy	1196 GCGCAGAGTGTGTGTGGCGCTGCTCTCGGCGGTGTGGCGCGCGCTGTGACAGCAGCGCAG	1255		
Db	481 GCGCAGAGTGTGTGTGGCGCTGCTCTCGGCGGTGTGGCGCGCGCTGTGACAGCAGCGCAG	540		
Qy	1256 CTTGGGCTCTCAGAGGCGGAGAGGCGACTTTCAGCTTCGCGACCTGTGGGAGCTGTGAGCGCA	1315		
Db	541 CTTGGGCTCTCAGAGGCGGAGAGGCGACTTTCAGCTTCGCGACCTGTGGGAGCTGTGAGCGCA	600		
Qy	1316 GGTCTTTCAGGTTGGGCGGAGATGATCGACCAACATGAGATGAAAGTCAACGTGCGCGCTG	1375		
Db	601 GGTCTTTCAGGTTGGGCGGAGATGATCGACCAACATGAGATGAAAGTCAACGTGCGCGCTG	660		
Qy	1376 GACCGTGCAGAGCCCGGCGAGGCGGCGCGGCGCGGACTCTGTTCACGCTACGCGCGCGCC	1435		
Db	661 GACCGTGCAGAGCCCGGCGAGGCGGCGCGGCGCGGACTCTGTTCACGCTACGCGCGCGCC	720		
Qy	1436 CTCCTCGGTCGTGTCTTTCAGAGAGCGGCGGCGGCTTTCGACCCCGAGAGAGCGCTGTGC	1495		
Db	721 CTCCTCGGTCGTGTCTTTCAGAGAGCGGCGGCGGCTTTCGACCCCGAGAGAGCGCTGTGC	780		
Qy	1496 CGCCATCTTTCGCGCGCGGCTGTGTGGCGGCTGTGTGCGCTTTCAGCCGCTGTGTGCGAA	1555		
Db	781 CGCCATCTTTCGCGCGCGGCTGTGTGGCGGCTGTGTGCGCTTTCAGCCGCTGTGTGCGAA	840		
Qy	1556 GCTGAGCTGA 1565			
Db	841 GCTGAGCTGA 850			

	Matches	852;	Conservative	0;	Mismatches	209;	Indels	0;	Gaps	0		
OY	495	GCTTGGCCCGGAGTTGGGCA	CCCA	CGAGAGATGGGGAC	CCGCA	CCCTCA	CTTCGACGGAG	554				
Db	1	GTATTGTTTCGAGATTGGATT	TTTA	TACGGAGATGGGGA	TGTTT	TTTAA	GTTCGTAAGGAG	60				
OY	555	CCA	CCGTGAGAGCCAGAGCCG	TGCGAGAGACA	CGA	CGTGTGA	CTCGAGTGGCC	CTGGAGGA	614			
Db	61	TTATCTGTGAGAGGTTA	AGGG	CCGTGTAAAGATA	TACGA	CGTGTAT	TGGAGATGGCGTT	TGGGGA	120			
OY	615	GGATGGA	CGAGGGAGCCGGGGGA	CCGCTTA	CCGGGCTCC	CTGGCGCC	CCCGTCCGCA	674				
Db	121	GGATGGA	CGAGGGAGCCGGGGAT	GTTTAA	CCGGGTTTTT	TTTTCGCGCTT	CGTTCGTAGA	180				
OY	675	GGGCA	GTCGAGAGGATCCCGGG	CGGGGCTCCG	TGGA	CGTTGG	CGGTAAGCGCGGAGCGATC	734				
Db	181	GGGCTATGTCGAGGGATTT	CGGACGGATTT	CGTGA	CGTTGG	CGGTAAGCGCGGATTC	CGAGGATTT	240				
OY	735	ACGGACCATGA	GAAGCGTTT	CGTGC	CGCGCGG	CCCA	AGGCGGGATTA	GGCCAT	794			
Db	241	ACGGATTTATGA	GAAGCGTTT	CGTGC	CGCGGTTT	AA	GGGTACGGATTA	GGGGTTA	300			
OY	795	CTTGC	CGCGCTGAAGGGGAGGCTTA	CGGGCGCGGGCGCGCGG	CCCA	CGCCGAG	CGCCAC	854				
Db	301	TTTGTCCGTTGA	AGGAGGATTA	CGGACCGG	CGCGCTCG	GGTTTAA	CTGAA	974				
OY	855	GCGATGG	CGAGGAGAGAGTGC	CAAGGCGCTG	TCGA	CGGGCTCA	CAACAGCA	CTGCCTATC	914			
Db	361	GCGATGG	CGAGGAGAGAGTGA	AGCGGTTT	TGGA	CGGGTTTAA	TAAACGATTC	CGTGT	420			
OY	915	TACCA	CACTGGTGTCTGA	CCGTGTGGTGT	CGCTCG	CGGACTG	CGACAA	CTTGGCGAGAG	974			
Db	421	TATTATTATTTGGTGTGAT	CTGTGTGGTGT	CGCGGATTC	GTGAA	TTTTCGGTAA	GGAG	480				
OY	975	CTGCA	AAAGA	CGCGCCAGAGGCGCGAGAG	CTGGCGGGTGT	CCACTGTG	CGCCGCTGACT	1034				
Db	481	TTGTAA	AAAGACGCTTAA	AGGCGCTTA	AGGATTTG	CGGTTTAA	TTTGGGTT	CGTTGAT	540			
OY	1035	GCTGTG	CTGCGGACCGGGGCGCTTG	ACCGCGCGGCGCGG	CCGAGTTT	GACGCGCTCTGG	1094					
Db	541	GTGTGTGTTGGCGGAT	TCGGGGTTTGTGT	CGTCA	CGAGCGCTG	CGAGTTTGA	CGGTTTGG	600				
OY	1095	GTGGCC	TTCTCGGGCTGCGCTGGA	CACTGTGGAA	CGGACA	NGCGACGCTG	CGTGGAGCTG	1154				
Db	601	GTGGTTTTTTCGGGTTT	GTGTGGA	ATTTGTGGA	AGCGATTA	NGCGACGCGCTT	GGAATTTG	660				
OY	1155	GGCG	CGCGCTTCCCGCTG	CA	CGCGCGCGCGGCA	CCGCTG	GTGCGCA	CGGTGTGCG	1214			
Db	661	GGCGT	CGCGTTTTTCTGTGTA	CGCTCGCGCGGCTG	TGAGCGTA	TAGGTGTG	GTGGC	720				
OY	1215	GCC	CTCCCGCGGTGGGCGCGCGCGCTG	AGCA	CCGGA	CCCGAC	CTTGGCGTGA	GGCGGAG	1274			
Db	721	GTTTTTTTCGGCGGTGGCGCGCGCGCTG	AGATTCGTA	TTTCGTA	TTTTCG	GGTTTTCGAGCGGAG	780					
OY	1275	GGCA	CTTTCGACCTGCGGGA	CCCTGCGGAG	CTGGA	CGCGGATCTT	CA	GTGGCGAG	1334			
Db	781	GGCGATTTTCGACG	TCGCGGATTTTCG	GGGAGTTTGA	AGCGCA	GGTTTTTAA	GGTGGCGAG	840				
OY	1335	ATGAT	CGACAA	CATGGA	AGTAA	AGTCA	CGTCCCGCTGGA	CCGTGCA	ACC	CGGCA	1394	
Db	841	ATGAT	TCATATA	TATGGA	GATGA	AGTAA	CGTGTTCGTTG	ATCGTGA	TTTCGATG	900		
OY	1395	GGCG	CGGCGCGCGAC	TCGTCCAC	CGGTAC	AGCGCGG	CCCTCT	CGTGTG	CGTGT	1454		
Db	901	GGCGCGGCGCGT	CGAGTTTTTGT	TTTACG	GTTA	AGCGTCCG	TTTTTTTT	CGGTGTG	GTGTTTGG	960		
OY	1455	CAGAG	CGCGGCGGCGGTTT	CGACCCCA	CGA	AGGCGCTT	GC	CGCATC	CTTTTCG	CGCC	1514	
Db	961	TAGAG	CGCGGCGGCGGTTT	CGATTTTAA	GGA	AGGTTTGTG	TCGTTAT	TTTTTTTCG	CGCTC	1020		
OY	1515	GTGCTG	CTGGCGGCTGTG	CGCTTAC	GCGGTG	CGGTGGCGGA	1555					
Db	1021	GTGTTGT	TGGCGGTTGTGTTTTT	AGTGTG	GTGCGG	GGA	1061					





```

Db 241 ACGGATTAAGAGAGGCTTCTGTCTGCGCGGCTTTAAGTCGCGATGGGGTTATAT 300
Qy 795 CCTGCCGCTGAGAGGGAGGCTTAACGCGCGCGCGCGCGCGCCAGCCGAGCCACC 854
Db 301 TTGTGCGCGTTAGAGGGGAGGTTAAGCGCGCGCGCGCGCGGTTAGTCGAGTTATC 360
Qy 855 GCGATGGCGAGAGAGAGTGCAGAGGCGCTGTGACCGCGCTCAACAGACGCTGCTGC 914
Db 361 GCGATGGCGAGAGAGTGCAGAGGCGCTGTGACCGCGCTTTATATAGACATTCGCT 420
Qy 915 TACCAACACCTGTGCTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 974
Db 421 TATTAATTAATTTGGTGTGATCGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
Qy 975 CTGCAAAAGACGCGCCAGAGAGCGCGACGAGCTGCGCTGCTCACTGCGCCGCTG 1034
Db 481 TTGTAAAGACGCGTTAGAGAGCGGTGAGAGTTGCGCTTTATTTGCGCTTGATT 540
Qy 1035 GCTGTGCTGCGGACCGCGGCGCTGCGCGCGACGAGCGCGCGAGTTGAGCGCTG 1094
Db 541 GTTGCTGTGCGGATGCGGAGTTGCTGCTGACGAGCGCGCTGAGTTGAGCGGTT 600
Qy 1095 GTGACCTTCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1154
Db 601 GTGTGTTTTTTCGCGTTGTTGATTTGTTGAAAGCGGATATGCGACGCGCGTTG 660
Qy 1155 GCGCGCGCGTTCCCGCTGACGCGCGCGCGCGACGCGCTGCTGCTGCTGCTGCT 1214
Db 661 GCGGTGCGCTTTCTTTGTAACGCGCTGCGCGCGCTGCTGCTGCTGCTGCTGCT 720
Qy 1215 GCTCTCTCGCGGCTGCGGCGCGCGCTGAGCACCGCGACCTGCGGCTGCGAGCG 1274
Db 721 GTTTTTTTCGCGGCTGCGGCGCGCGCTGAGTATTCGATTTGCGTTTCAAGCGAG 780
Qy 1275 GCGGACTTTCGAGCTGCGGACCTGCGGAGCTGAGCGCGCGAGGCTCTTCAAGT 1334
Db 781 GCGGATTTTCGAGCTGCGGATTTTGGCGGAGTTGAGACGCGAGGTTTTTAAAGT 840
Qy 1335 ATGATGACACATGAGATGAAAGTCAACGTCGCCCGCTGACCGTCAAGCCGCG 1394
Db 841 ATGATGATTAATATGAGATGAAAGTCAACGTCGTTCTGTTGATTCGTAAGTT 900
Qy 1395 GCGGCGGCGCGGAGCTCCGTCGACGCTGACGCGCGCGCGCTCCGCTGCTGCT 1454
Db 901 GCGGCGGCGCTGCAAGTTTGTGTTACGTTTACGCTGCTGCTTTCGCTGCTG 960
Qy 1455 CAGGACGCGGCGGCGGCTGCGACCCGAGAGAGCCCTGCGCGCATCTTTCGCG 1514
Db 961 TAGGACGCGGCGGCGGCTGCGATTTTAAAGAAAGTTTGGTGTATTTTTCGCG 1020
Qy 1515 GTGCTGCTGCGGCTGCTGCGCCCTGACCGCTGCTGCTGCTGCGCGAG 1555
Db 1021 GTGTTGTGCGGCTGTGTGTTTATGTCGTGTGCGTGGCGAA 1061

```

RESULT 5  
US-10-363-483a-27372/c

```

; Sequence 27372, Application US/10363483A
; Publication No. US20050064401A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Beilin
; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain
; FILE REFERENCE: 82011
; CURRENT APPLICATION NUMBER: US/10/363,483A
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 27372
; LENGTH: 1061
; TYPE: DNA

```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-Island No: 27372
US-10-363-483a-27372

Query Match      23.3%; Score 726.6; DB 9; Length 1061;
Best Local Similarity 80.3%; Pred. No. 2,1e-174;
Matches 852; Conservative 0; Mismatches 209; Indels 0; Gaps 0;

Qy 495 GCTTCCCGGAGTTTGACCCACGAGAGATGGGACCGGACCTTCACTTCCAGCGAG 554
Db 1061 GTTGTTCGAGAGTTGATTTACGAGAGATGGGATGATTTTATTTAGTTTCGTAGG 1002
Qy 555 CCAACCTTGGAGGCGGCGGCTGTCAGACACAGATGTGACTGGAGTGGCCCTGGGGA 614
Db 1001 TTATGTGAGAGTTTGGCGCGGTGTGAGATACGAGTGTGATTCGAGGTGCTTTGG 942
Qy 615 GGAATGACGAGGAGACCGGGGACCGCTTAAACGGGGCTCCCTCTGCGCGCCGTCG 674
Db 941 GGAATGACGAGGAGACCGGGGATCGTTTAAACGGGGTTTTTTTTCGCGCTTTCGT 882
Qy 675 GCGGACGTCGAGAGGCTCCCGGCGGCTCCGTTGACGTTTGGCGTTAGCGCGAGTC 734
Db 881 GCGGTACGTCGAGAGGCTTCGCGCGGCTTCGTTGACGTTTGGCGGTAGCGTCA 822
Qy 735 ACGGACATTAAGAGACGTTTCGTCGCGCGCGCCAGAGCGGAGTTAGCCACAT 794
Db 821 ACGGATTAAGAGACGTTTCGTCGCGCGGTTTAAAGTCGAGATGGAGGTTAGTTAT 762
Qy 795 CTTGCGGCTGAGAGGAGAGCTTAAACGAGCGCGCGCGCGCGCCAGCCGAGCCAC 854
Db 761 TTTGTGCGGTTAGAGGAGAGGTTTAAACGCGCGCGCGCGGCTTAACTGAGTTAT 702
Qy 855 GCGATGGCGAGAGAGAGTCAAGGCTGCTGACGCGGCTCAACAAACGACTGCTGC 914
Db 701 GCGATGGCGAGAGAGAGTCAAGGCTGCTGACGCGGCTTAAACAAACGACTGCTGC 642
Qy 915 TACCAACACCTGCTGACCGTTCGTTGCTGCGGAGCTGCGAAGCTGCGGAGAG 974
Db 641 TATTAATTAATTTGCTGATTCGTCGCTGCTGCGGAGTTGCTAATTTGCGTTGAT 582
Qy 975 CTGCAAAAGACGCGCAGAGAGCGAGAGCTGCGGCTGCTCAACCTGCGCGCTGACT 1034
Db 581 TTGTAAGAACCGCTTAAAGAGCGTAAAGAGTTGGCGGTTTATTTGCGTTGAT 522
Qy 1035 GCTGTGCTGCGGACCGGCGGCTTGCACCGCGCGACGCGCGAGTTGAGCGCTG 1094
Db 521 GTTGTTGCTGCGGATCGGCGGTTGCTGCTGACGAGCGGCTGAGGTTGAGG 462
Qy 1095 GTGCGCTTCTCGGCGCTGACCTGCTGAGAGGAGCAATGCGAGCTGCTGAGAGCT 1154
Db 461 GTGTGTTTTTTCGAGGTTGTTGATTTGTTGAGAGGAGATATGCGACGCGCTTGA 402
Qy 1155 GCGCGCGGCTTCCCGCTGACCGCGCGCGCGACCGCTGTGTCGACAGGTGCTG 1214
Db 401 GCGGTGCGGTTTTCTGTTGATCGCGTGCAGCGCGCTGTTGTCGTATAGTGTG 342
Qy 1215 GCTCTCTCGCGGCTGCGGCGCGCGCTGAGACCCGCGAGCTGCGCTGAGCGAG 1274
Db 341 GTTTTTTTCGCGGCTGCGGCGCGCGCTTGAATATTCGATTTTCGCTTCAAGCG 282
Qy 1275 GCGGACTTTCGAGTGGCGGACCTGCGGAGAGCTGAGAGCGAGAGTCTTCAAGTGG 1334
Db 281 GCGGATTTTCGAGTGGCGGATTTGCGGAGGTTGAGAGCGAGAGTTTAAAGTGG 222
Qy 1335 ATGATGACACATGAGATGAAAGTCAACGTCGCCCGCTGACCGTCAAGCCGCG 1394
Db 221 ATGATGATTAATATGAGATGAAAGTCAACGTTTTCGTTGATGCTGTAAGTTCG 162
Qy 1395 GCGGCGGCGCGGAGCTCTGTCACAGCTGACGCGCGCGCTCTCTGCTGCTGCT 1454
Db 161 GCGGCGGCGCTGAGATTTTGTTTAAGCTTACGCTGCTGCTTTCGCTGCTGCT 102

```

QY 1455 CAGAGCGGGGGGGTTCAGACCCAGAGAGCCCTGAGCCGACATCTTTTTCGGCGC 1514  
DB 101 TAGAGCGGGGGGGGGTTCAGATTTTAGAGATTTTGTCTTATTTTTCGGCGC 42  
QY 1515 GTGCTGCTGAGCGGCTGTGACCTTACCGCTGTGCTGAGCGCA 1555  
DB 41 GTGTTGTGCGGGTGTGTGTTTGTAGTGTGTGCTGAGCGCA 1

RESULT 6  
US-10-029-386-20699/c  
Sequence 20699, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
FILE REFERENCE: AEOMICA-X-2  
CURRENT APPLICATION NUMBER: US/10/029,386  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 20699  
LENGTH: 708  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO AC010615.5  
OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.56  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3  
OTHER INFORMATION: EST HUMAN HIT: AW302149.1, EVALUR 0.00e+00  
OTHER INFORMATION: SWISSPROT HIT: P31735, EVALUR 5.50e+00  
US-10-029-386-20699

Query Match 22.6%; Score 706.4; DB 6; Length 708;  
Best Local Similarity 99.9%; Pred. No. 2.6e-169;  
Matches 707; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 858 ATGGCGAGGAGAGAGTGCAGAGCGCTGTGACGGGCTCAACAAGACGACTGCTTAC 917  
DB 708 ATGGCGAGGAGAGAGTGCAGAGCGCTGTGACGGGCTCAACAAGACGACTGCTTAC 649  
QY 918 CACCACTGTGTGACCGTGTGCTGTGCGGAGCTCGACAACTGTGAGAGAGAGCTG 977  
DB 648 CACCACTGTGTGACCGTGTGCTGTGCGGAGCTCGACAACTGTGAGAGAGAGCTG 589  
QY 978 CAAAAGACGCGCAGAGAGCGCAGAGCTGTGTGTCACTGTGCGCCCGCTGACTGCT 1037  
DB 588 CAAAAGACGCGCAGAGAGCGCAGAGCTGTGTGTCACTGTGCGCCCGCTGACTGCT 529  
QY 1038 GTGCTGCGGAGCGGGGCTGTGCGCGCAGAGCGCGGAGTTGAGAGGAGCTTGGT 1097  
DB 528 GTGCTGCGGAGCGGGGCTGTGCGCGCAGAGCGCGGAGTTGAGAGGAGCTTGGT 469  
QY 1098 GCGTTCTGGGAGCTGTGACCTGTGAGAGAGGAGCATGAGACGCTGCTGAGAGCTGG 1157  
DB 468 GCGTTCTGGGAGCTGTGACCTGTGAGAGAGGAGCATGAGACGCTGCTGAGAGCTGG 409  
QY 1158 GCGCGCTTCCGCTGTGACCGCGCGCGAGACCGCTGTGTGCGCAAGATGTGTGCGCG 1217  
DB 408 GCGCGCTTCCGCTGTGACCGCGCGCGAGACCGCTGTGTGCGCAAGATGTGTGCGCG 349  
QY 1218 TCTCTCGGAGTGTGCGCGCGCGCGCTGAGACCTCGAGCTTGGGCTCGAGGCGAGG 1277  
DB 348 TCTCTCGGAGTGTGCGCGCGCGCGCTGAGACCTCGAGCTTGGGCTCGAGGCGAGG 289

QY 1278 GACTTGAAGTCGCGAGACCTGCGGAGAGCTGAGACCGAGAGTCTTTCAGGTGGCGAGATG 1337  
DB 288 GACTTGAAGTCGCGGAGACCTGCGGAGAGCTGAGACCGAGAGTCTTTCAGGTGGCGAGATG 229  
QY 1338 ATGACAAATGAGATGAGATGAGTCAACGTGCCCCCTGTGACCTGTGACCGCGAGCG 1397  
DB 228 ATGACAAATGAGATGAGATGAGTCAACGTGCCCCCTGTGACCTGTGACCGCGAGCG 169  
QY 1398 GCGGCGCGGAGCTCTGTGCAACGAGTACGCGCGCGCGCCCTCTCGGTGTGTCTTGCAG 1457  
DB 168 GCGGCGCGGAGCTCTGTGCAACGAGTACGCGCGCGCGCCCTCTCGGTGTGTCTTGCAG 109  
QY 1458 GAGCGCGGCGGAGTGTGCGACCCAGAGAGCGCTGTGCGCGCATCTTTCGCGCGCTG 1517  
DB 108 GAGCGCGGCGGAGTGTGCGACCCAGAGAGCGCTGTGCGCGCATCTTTCGCGCGCTG 49  
QY 1518 CTGCTGCGGCTGTGAGCCCTTACCGCTGTGCGGTGCGAGAGCTGAGCTGA 1565  
DB 48 CTGCTGCGGCTGTGAGCCCTTACCGCTGTGCGGTGCGAGAGCTGAGCTGA 1

RESULT 7  
US-10-363-345A-27369/c  
Sequence 27369, Application US/10363345A  
Publication No. US20040234960A1  
GENERAL INFORMATION:  
APPLICANT: Alexander Olek  
APPLICANT: Christian Piepenbrock  
APPLICANT: Kurt Berlin  
TITLE OF INVENTION: Method for determining the degree of methylation of defined  
FILE REFERENCE: E01/1227  
CURRENT APPLICATION NUMBER: US/10/363,345A  
NUMBER OF SEQ ID NOS: 40712  
SEQ ID NO 27369  
LENGTH: 1061  
TYPE: DNA  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
OTHER INFORMATION: CpG-island No: 27369  
US-10-363-345A-27369

Query Match 19.0%; Score 594.4; DB 8; Length 1061;  
Best Local Similarity 72.5%; Pred. No. 9.8e-141;  
Matches 769; Conservative 0; Mismatches 291; Indels 0; Gaps 0;

QY 496 CTGGCCCGGAGTGTGCAACCCAGAGAGATGGAGACCGCACTTCAGCTTTCGAGGAGAC 555  
DB 1060 CTGGCCCGGAGTGTGCAACCCAGAGAGATGGAGACCGCACTTCAGCTTTCGAGGAGAC 1001  
QY 556 CACCGTGAAGCGCAGGCGGTGACAGACACAGCTGTGACTGTGAGATGCGCTGGGAG 615  
DB 1000 CACCGTGAAGCGCAGGCGGTGACAGACACAGCTGTGACTGTGAGATGCGCTGGGAG 941  
QY 616 GATGACGAGGAGAGCGGAGACCGCTTAACGGGAGCTTCTGTGCGCGCCCGCTGCGCAG 675  
DB 940 GATGACGAGGAGAGCGGAGACCGCTTAACGGGAGCTTCTGTGCGCGCCCGCTGCGCAG 881  
QY 676 GCGCAGTGTGAGAGTCCCGGCGGAGCTCGGTGAGAGTGTGCGGTGAGCGCCGAGAGAGTCA 735  
DB 880 GCGCAGTGTGAGAGTCCCGGCGGAGCTCGGTGAGAGTGTGCGGTGAGCGCCGAGAGAGTCA 821  
QY 736 GCGACCATGAGAGCGTTCGTGCGCGCGCGCCCAAGCGCGGAGTGGGGTTAGCAATC 795  
DB 820 GCGACCATGAGAGCGTTCGTGCGCGCGCGCCCAAGCGCGGAGTGGGGTTAGCAATC 761  
QY 796 CTGCGCGCTGAGGAGGAGCTTAACGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 855  
DB 760 CTGCGCGCTGAGGAGGAGCTTAACGGCGCGGCGCGCGCGCGCGCGCGCGCGCGCGCG 701  
QY 856 CGATGCGAGAGGAGAGTGAAGCGCTGTGAGAGCGGCTCAACAAGACGACTGCTGTCT 915

```

Db      700 CCAATACGAAAAAAATACAAAAGCTACTTAACGAACTCAACAAAACGACTGACT 641
Qy      916 ACCACCACTGAGTGTGACCGTGTGAGTGTGAGCGAGACTTGCAGAACTTGGCGAGAGC 975
Db      640 ACCACCACTGAACTTAACCGTGTGATTAACGAGAACTTGCAGAAAACCTACGACAAAAC 581
Qy      976 TGCAGAAAGCGGCGAGAGAGGCGAGAGAGTGGCGGTGTCCACTGGCGCGCGCTGACG 1035
Db      580 TACAAAAACGCGCAAAAAACGCAAAAACTAACGATATCACTACGCCCACTAACCTAA 521
Qy      1036 CTGTGTGCGGCGACCGGAGCGCTGTGCGCGAGCGAGCGCGCGAGTTCAGAGCGGCTGTGG 1095
Db      520 CTATCTAGAGCGGACCGGAAACCTAACCGCGAGAGAGCGCGCGAATTCAGAACGACTTAA 461
Qy      1096 TGGCTCTTCTGAGGCTGCTGAGACTGTGTGAAAGCGGACATGAGCGCTGCTGAGAGCTGG 1155
Db      460 TAACTCTTCTGAACTTAACCTTAACCTTAACGAAACATAGAGCGCGCTTAAACTTA 401
Qy      1156 GGGCGCGGTTCCGCGTGAAGCGGCGGCGGAGCGCGTGTGAGCAAGGTGTGCTGGCG 1215
Db      400 AGCGCGGTTCCGCGTGAAGCGGCGGCGGAGCGGCTTAATAGCAATAATTAATTAACG 341
Qy      1216 CCTCTCTCGGCGTGTGCGGCGGCGGCTGTGAGCAACCGGAGCTGTGAGCGGAGG 1275
Db      340 CCTCTCTCGGCGTGAAGCGGCGGCGCTTAACACCGGAGACTTACGAAAGGAA 281
Qy      1276 GCGACTTGAAGCTGTGAGAGCTGTGAGAGCGGAGCGGAGTCTTTCAGGTGGCGGAG 1335
Db      280 ACGACTTGAAGCTGTGAGAGCTGTGAGAGCGGAGAGCTTTCAGGTGGCGGAG 221
Qy      1336 TGAATGACAACTGAGATGAGAGTCAAGTGTGCGGCGGCTGTGAGAGCGGAGG 1395
Db      220 TAAATGACAACTGAGATGAGAGTCAAGTGTGCGGCGGCTGTGAGAGCGGAGG 161
Qy      1396 CGGCGGCGGCGGAGCTGTGAGAGCGGCTGTGAGAGCGGCGGCTGTGAGAGCTGTGAG 1455
Db      160 CGAGCGAGCGGCGGAGCTGTGAGAGCGGCTGTGAGAGCGGCGGCTGTGAGAGCTGTGAG 101
Qy      1456 AGGAGCGGCGGAGGAGTGTGAGAGCGGCGGAGCGGCGGCTGTGAGAGCTGTGAGAG 1515
Db      100 AAAAAAGCGGAGAGAGTGTGAGAGCGGCGGAGCGGCGGCTGTGAGAGCTGTGAGAG 41
Qy      1516 TGTGTGTGCGGCGGCTGTGAGAGCGGCGGCTGTGAGAGCGGCGGCTGTGAGAG 1555
Db      40 TACTACTTAACGACTTAACCTTAACCGTATACGTAAAGAA 1

RESULT 8
US-10-363-345A-27370
; Sequence 27370, Application US/10363345A
; Publication No. US20040234960A1
; GENERAL INFORMATION:
; APPLICANT: Alexander Olek
; APPLICANT: Christian Piepenbrock
; APPLICANT: Kurt Berlin
; TITLE OF INVENTION: Method for determining the degree of methylation of defined
; FILE REFERENCE: E01/1227
; CURRENT FILING DATE: 2003-03-03
; NUMBER OF SEQ ID NOS: 40712
; SEQ ID NO 27370
; LENGTH: 1061
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; OTHER INFORMATION: CpG-island No: 27370
US-10-363-345A-27370
Query Match 19.0%; Score 594.4; DB 8; Length 1061;
Best Local Similarity 72.5%; Pred. No. 9.8e-141;

```

```

Matches 769; Conservative 0; Mismatches 291; Indels 0; Gaps 0;
Qy      496 CTGGCCCGGAGGTTGGCAACCCAGAGAGATGGGAGACCGCAACCTTACGCTTGGCAGGAGC 555
Db      2 CTTACCCCGGAAATTAACCCAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATGAGAG 61
Qy      556 CACCGTGAAGGCGAGGCGGTGAGAGACAGAGCGTGTGACTGTGAGAGTGTGAGAGTGTGAG 615
Db      62 CACCGTGAAGGCGAGGCGGTGAGAGACAGAGCGTGTGACTGTGAGAGTGTGAGAGTGTGAG 121
Qy      616 GATGAGCAGAGGAGCGGAGGAGCGCTTAACGAGGCTTCCCTGTGAGCGGCGGCTGTGAGAG 675
Db      122 AATAAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 181
Qy      676 GGGCACTGTGAGAGGTTCCGCGGCGGCTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 735
Db      182 ACGCACTGTGAGAGGTTCCGCGGCGGCTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAG 241
Qy      736 CCGACCATGAAAGAGGTTCTGTGCGCGGCGGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAG 795
Db      242 CGAACCATTAAGAGGTTCTGTGCGCGGCGGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAG 301
Qy      796 CTGCGCGCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 855
Db      302 CTACCGCGCTTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 361
Qy      856 CGATGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 915
Db      362 CGATGAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 421
Qy      916 ACCACCACTGAGTGTGAGAGCGGCTGTGAGAGCGGAGCTGTGAGAGCGTGTGAGAGCGAGAGC 975
Db      422 ACCACCACTTAATTAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAAC 481
Qy      976 TGCAGAAAGCGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1035
Db      482 TACAGAAAGCGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 541
Qy      1036 CTGTGTGTGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1095
Db      542 CTATCTCTGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 601
Qy      1096 TGGCTCTTCTGAGGCTGTGAGAGCTGTGAGAGCGGAGCGGAGCTGTGAGAGCTGTGAGAG 1155
Db      602 TAACTTCTGAGAGCTTAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAACCTTAAC 661
Qy      1156 GCGCGCGTTCGCGTGAAGCGGCGGCGGAGCGGCTGTGAGGAGCAGAGTGTGAGCGG 1215
Db      662 ACGCGCGTTCGCGTGAAGCGGCGGCGGAGCGGCTTAATTAACGAAATTAATTAACG 721
Qy      1216 CCTCTCTCGGCGTGTGCGGCGGCGGCTGTGAGAGCGGAGCGGCTGTGAGAGCGGAGG 1275
Db      722 CCTCTCTCGAGGTAAGAGCGGCGGCTTAACCGGAGCGGAGCTTAAGAGCTTAAGAGGAGG 781
Qy      1276 GCGACTTGAAGCTGTGAGAGCTGTGAGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1335
Db      782 ACGACTTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 841
Qy      1336 TGAATGACAACTGAGATGAGAGTGTGAGAGCGGCGGCTGTGAGAGCGGAGGAGGAGGAG 1395
Db      842 TAAATGACAACTGAGATGAGAGTGTGAGAGCGGCGGCTTAACCGGAGCGGAGGAGGAGGAG 901
Qy      1396 GGGCGGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1455
Db      902 CGAGCGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 961
Qy      1456 AGGAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1515
Db      962 AAAAAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1021
Qy      1516 TGTGTGTGCGGCGGCTGTGAGAGCGGCGGCTGTGAGAGCGGCGGCTGTGAGAGGAGGAG 1555
Db      1022 TACTACTTAACGACTTAACCTTAACCGTATACGTAAAGAA 1061

```

RESULT 9  
 US-10-363-483A-27369/C  
 ; Sequence 27369, Application US/10363483A  
 ; Publication No. US20050064401A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Alexander Olek  
 ; APPLICANT: Christian Piepenbrock  
 ; APPLICANT: Kurt Berlin  
 ; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain  
 ; TITLE OF INVENTION: Illnesses  
 ; FILE REFERENCE: 82011  
 ; CURRENT APPLICATION NUMBER: US/10/363,483A  
 ; CURRENT FILING DATE: 2003-03-03  
 ; NUMBER OF SEQ ID NOS: 40712  
 ; SEQ ID NO 27369  
 ; LENGTH: 1061  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
 ; OTHER INFORMATION: Cpg-Island No: 27369  
 US-10-363-483A-27369

Query Match 19.0%; Score 594.4; DB 9; Length 1061;  
 Best Local Similarity 72.5%; Pred. No. 9,8e-141;  
 Matches 769; Conservative 0; Mismatches 291; Indels 0; Gaps 0;

QY 496 CTTGCCCCGAGATTGGCACCCACGAGATGGGAGCCGACCTTCAGTTCCGAGGAGC 555  
 |||  
 DB 1060 CTTACCCCGAATTATACCCGAGAAATTAATAACCGACCTTCAGTTCCGAGAAAC 1001  
 |||  
 QY 556 CACCGTGAAGGCCAGGGCGGTGACAGACACGAGTGTGATCTGAGAGTGGCGTGGAG 615  
 |||  
 DB 1000 CACCGTGAAGGCCAGGGCGGTGACAGACACGAGTGTGATCTGAGAGTGGCGTGGAG 941  
 |||  
 QY 616 GATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 675  
 |||  
 DB 940 AATTAACGAGAAACGAGAAACGAGTGAAGAACTCCCTCTACGCGCCCGCTCCGAGAA 881  
 |||  
 QY 676 GCGCAGTGTGAGGAGTCCCGGCGGAGCTCCGTGAGACGTGGCGGTGACGCGGAGCGATCA 735  
 |||  
 DB 880 AGCGACGTGAGAAATCCCGAGCAACTCCGTAAACCTTAACATTAACGCGGAGCGATCA 821  
 |||  
 QY 736 CGGACCATGAAGAGCGTTGTCGTCGCGCGCGCCGAGCGCGGAGTGGGGTTAGCCATATC 795  
 |||  
 DB 820 CGAACCATTAATAACGTTGTAACGCGCGGAGCCGAGAAATTAATAACGATATC 761  
 |||  
 QY 796 CTTGCGCGGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 855  
 |||  
 DB 760 CTACCGCGGTAAATAATAATACTTAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 701  
 |||  
 QY 856 CGATGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 915  
 |||  
 DB 700 CGATTAACGAGAAATAATAATACTTAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 641  
 |||  
 QY 916 ACCACCACTGTGTGCTGACCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 975  
 |||  
 DB 640 ACCACCACTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 581  
 |||  
 QY 976 TGGCAAAAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1035  
 |||  
 DB 580 TACAAATAAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 521  
 |||  
 QY 1036 CTGTGTGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1095  
 |||  
 DB 520 CTATTAATTAACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 461  
 |||  
 QY 1096 TGGCTTTCTGGGCTGCTGAGACCTGCTGAGAGCGGACATGGAAGCGCTGCTGAGAGCTGG 1155  
 |||  
 DB 460 TAACCTTTCTGGAATCACTTAACCTTAATAAAGAGGAGGAGGAGGAGGAGGAGGAGGAG 401  
 |||

QY 1156 GCGCGCGTTCCTCGTGAACGCGCGGCGGAGCCGCTGGTGGCAAGGTGTGCTGCGG 1215  
 |||  
 DB 400 AGCGCGGTTCCTCGTGAACGCGCGGCGGAGCCGCTAATAGCAATATATATATACGAG 341  
 |||  
 QY 1216 CTTCTTCGCGGCTGGCGGCGGCGGCTGAGACCTCGAGCTTCGAGGCGGAGG 1275  
 |||  
 DB 340 CTTCTTCGAGCGTAAACGAGCGCGGCTTAACACCGGCAACCTTACGACTCGAAAGGAAA 281  
 |||  
 QY 1276 GCGACTTGAAGTTCGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1335  
 |||  
 DB 280 AGGACTTGAAGTTCGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 221  
 |||  
 QY 1336 TGAATGACAACTGAGAGTGAAGTCAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAG 1395  
 |||  
 DB 220 TGAATGACAACTAATAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 161  
 |||  
 QY 1396 GCGCGGCGGCGGAGCTCTGTTCACGCTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1455  
 |||  
 DB 160 CGACGAGCGCGAGACTCTGATCAAGATCAACGCGGAGCGGCGGCGGCGGCGGCGGCGG 101  
 |||  
 QY 1456 AGGAGCGGCGGCGGAGCTCTGTTCACGCTGACGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1515  
 |||  
 DB 100 AAAAAAGGAGAAATAATTAACGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 41  
 |||  
 QY 1516 TGGCTGTGCGGCTGTGAGGCGCTAAGCGGTGTGCGGAG 1555  
 |||  
 DB 40 TACTTACTAAGACTTAATTAACCTTAACGATATACGTAACGAA 1

RESULT 10  
 US-10-363-483A-27370  
 ; Sequence 27370, Application US/10363483A  
 ; Publication No. US20050064401A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Alexander Olek  
 ; APPLICANT: Christian Piepenbrock  
 ; APPLICANT: Kurt Berlin  
 ; TITLE OF INVENTION: Diagnosis of illnesses or predisposition to certain  
 ; TITLE OF INVENTION: Illnesses  
 ; FILE REFERENCE: 82011  
 ; CURRENT APPLICATION NUMBER: US/10/363,483A  
 ; CURRENT FILING DATE: 2003-03-03  
 ; NUMBER OF SEQ ID NOS: 40712  
 ; SEQ ID NO 27370  
 ; LENGTH: 1061  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
 ; OTHER INFORMATION: Cpg-Island No: 27370  
 US-10-363-483A-27370

Query Match 19.0%; Score 594.4; DB 9; Length 1061;  
 Best Local Similarity 72.5%; Pred. No. 9,8e-141;  
 Matches 769; Conservative 0; Mismatches 291; Indels 0; Gaps 0;

QY 496 CTTGCCCCGAGATTGGCACCCACGAGATGGGAGCCGACCTTCAGTTCCGAGGAGC 555  
 |||  
 DB 2 CTTACCCCGAATTATACCCGAGAAATTAATAACCGACCTTCAGTTCCGAGAAAC 61  
 |||  
 QY 556 CACCGTGAAGGCCAGGGCGGTGACAGACGAGTGTGATCTGAGAGTGGCGTGGAG 615  
 |||  
 DB 62 CACCGTGAAGGCCAGGGCGGTGACAGACGAGTGTGATCTGAGAGTGGCGTGGAG 121  
 |||  
 QY 616 GATGACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 675  
 |||  
 DB 122 AATTAACGAGAAACGAGAAACGAGTGAAGAACTCCCTTAACGAGGAGGAGGAGGAGGAG 181  
 |||  
 QY 676 GCGCAGTGTGAGGAGTCCCGGCGGAGCTCCGTGAGACGTGGCGGTGACGCGGAGCGAGTCA 735  
 |||  
 DB 182 AGCGACGTGTGAGAAATCCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 241  
 |||  
 QY 736 CGGACCATGAAGAGCGTTGTCGTCGCGCGCGCCAGGCGCGGAGTGGGGTTAGCCATATC 795  
 |||

Db 242 CGAACATATAAAAGCTTGTACCGGCGCAACCAAAACGGAATATAAATTAACCAATC 301  
Qy 796 CTGCGCGCTGAGGGGAGGCTTAACGGGCGCGGGCCGAGCCGAGCCCAACCG 855  
Db 302 CTACCGCGCTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 361  
Qy 856 CGATGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 915  
Db 362 CGATTAACGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 421  
Qy 916 ACCACACCTGAGTGTGACCTGCTGAGTGTGCTGAGGAGCTGAGGAGCTGAGGAG 975  
Db 422 ACCACACCTTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 481  
Qy 976 TGCAGAAAGAGCGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1035  
Db 482 TACAAAAAGAGCGGCGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 541  
Qy 1036 CTGTGCTGGGAGAGCGGGGCTGGGCGGAGAGGAGGAGGAGGAGGAGGAGGAGGAG 1095  
Db 542 CTATATACGAGCGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 601  
Qy 1096 TGGCTCTTCGAGGCTGCTGAGCTGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1155  
Db 602 TAACTCTTCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 661  
Qy 1156 GGGCGCGCTTCGCGCTGAGCGGCGGCGGAGCGGCTGCTGAGGAGGAGGAGGAGG 1215  
Db 662 AGCGCGCGCTTCGCGCTGAGCGGCGGCGGAGCGGCTGCTGAGGAGGAGGAGGAGG 721  
Qy 1216 CCGCTCCGCGGCTGGGCGGCGGCGGCTGAGGAGCGGCGGCTGAGGAGGAGGAGGAGG 1275  
Db 722 CCGCTCCGCGGCTGGGCGGCGGCGGCTGAGGAGCGGCGGCTGAGGAGGAGGAGGAGG 781  
Qy 1276 GCGACTTCGAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1335  
Db 782 AGGACTTCGAGCGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 841  
Qy 1336 TGAATGAGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1395  
Db 842 TGAATGAGCAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 901  
Qy 1396 GGGCGCGGCTTCGCGCTGAGCGGCGGCGGAGCGGCTGCTGAGGAGGAGGAGGAGG 1455  
Db 902 CGAGGAGCGCGGAGCTTCATTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 961  
Qy 1456 AGGAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1515  
Db 962 AAAAAAGCGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1021  
Qy 1516 TGTGCTGGGCGGCTGAGGCGGCTTACGCGGCTGAGGAGGAGGAGGAGGAGGAGG 1555  
Db 1022 TACTACTAACGACTATACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 1061

RESULT 11  
US-09-925-065A-740956/c  
; Sequence 740956, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925, 065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243, 096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252, 147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250, 092  
; PRIOR FILING DATE: 2000-11-30

; PRIOR APPLICATION NUMBER: US 60/261, 766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289, 846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 740956  
; LENGTH: 560  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-740956

Query Match 17.9% Score 559.6; DB 4; Length 560;  
Best Local Similarity 99.8%; Pred. No. 5.9e-132;  
Matches 559; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 199 CATCAAGCAAAAACTCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 258  
Db 560 CATCAAGCAAAAACTCCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 501  
Qy 259 CATCGTAAATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 318  
Db 500 CATCGTAAATTAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 441  
Qy 319 GCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 378  
Db 440 GCGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 381  
Qy 379 CCACTGATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 438  
Db 380 CCACTGATCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 321  
Qy 439 ACCCTACCGGCGCGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 498  
Db 320 ACCCTACCGGCGCGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 261  
Qy 499 GCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 558  
Db 260 GCCCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 201  
Qy 559 CGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 618  
Db 200 CGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 141  
Qy 619 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 678  
Db 140 GAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 81  
Qy 679 CAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 738  
Db 80 CAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 21  
Qy 739 ACCATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 758  
Db 20 ACCATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1

RESULT 12  
US-09-925-065A-769508/c  
; Sequence 769508, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925, 065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243, 096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252, 147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250, 092

```
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 769508
/ LENGTH: 557.
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-769508
```

```
Query Match      17.8%; Score 556.6; DB 4; Length 557;
Best Local Similarity 99.8%; Pred. No. 3.4e-131;
Matches 556; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 202 CACAGGCAAAAATCTCCGCGGAGCCTGCGCCGCTTTTACCTGAGCCTCAATTCCCAT 261
DB 557 CACAGGCAAAAATCTCCGCGGAGCCTGCGCCGCTTTTACCTGAGCCTCAATTCCCAT 498
QY 262 CCGTAATAATGAAACGGGTTGATCTCCGAGCGCTAACATTCCAGAACTCGATGGGCG 321
DB 497 CCGTAATAATGAAACGGGTTGATCTCCGAGCGCTAACATTCCAGAACTCGATGGGCG 438
QY 322 AAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 381
DB 437 AAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 378
QY 382 CTGATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 441
DB 377 CTGATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 318
QY 442 CTACCCGCGCCGCTTTGCGACGCGCTTAAGCGGAGCGCGGCTCTGCAAGCTTGCTTCC 501
DB 317 CTACCCGCGCCGCTTTGCGACGCGCTTAAGCGGAGCGCGGCTCTGCAAGCTTGCTTCC 258
QY 502 CCGGAGTTGGCACCACCGAGAGGATGGGAGCCGACCCCTGACGTTCCAGAGGAGCCACCT 561
DB 257 CCGGAGTTGGCACCACCGAGAGGATGGGAGCCGACCCCTGACGTTCCAGAGGAGCCACCT 198
QY 562 GAGAGCCAGGAGCGGTCAGAGACACGCTGTGACTCGAGATGCGGCTGCGGAGATGGA 621
DB 197 GAGAGCCAGGAGCGGTCAGAGACACGCTGTGACTCGAGATGCGGCTGCGGAGATGGA 138
QY 622 CGAGGAGCGGAGGAGCGGCTTAACGGGCTCCCTCTGCGCGCCCGCTCCGAGAGGCGAC 681
DB 137 CGAGGAGCGGAGGAGCGGCTTAACGGGCTCCCTCTGCGCGCCCGCTCCGAGAGGCGAC 78
QY 682 GTGAGGAGTCCCGGAGCGGCTCCGTGAGAGTGGCGGTAGCCGAGCGAGTCAAGAAC 741
DB 77 GTGAGGAGTCCCGGAGCGGCTCCGTGAGAGTGGCGGTAGCCGAGCGAGTCAAGAAC 18
QY 742 ATGAAGAGCGTTCTGTC 758
DB 17 ATGAAGAGCGTTCTGTC 1
```

## RESULT 13

```
US-09-925-065A-736351/c
/ Sequence 736351, Application US/09925065A
/ Publication No. US20050228172A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
```

```
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 736351
/ LENGTH: 556
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-736351
```

```
Query Match      17.8%; Score 555.6; DB 4; Length 556;
Best Local Similarity 99.8%; Pred. No. 6.1e-131;
Matches 555; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 203 ACAGGCAAAAATCTCCGCGGAGCCTGCGCCGCTTTTACCTGAGCCTCAATTCCCAT 262
DB 556 ACAGGCAAAAATCTCCGCGGAGCCTGCGCCGCTTTTACCTGAGCCTCAATTCCCAT 497
QY 263 CCGTAATAATGAAACGGGTTGATCTCCGAGCGCTAACATTCCAGAACTCGATGGGCG 322
DB 496 CCGTAATAATGAAACGGGTTGATCTCCGAGCGCTAACATTCCAGAACTCGATGGGCG 437
QY 323 AAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 382
DB 436 AAGGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 377
QY 383 TGATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 442
DB 376 TGATCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 317
QY 443 TACCCGCGCCGCTTTGCGACGCGCTTAAGCGGAGCGCGGCTCTGCAAGCTTGCTTCC 502
DB 316 TACCCGCGCCGCTTTGCGACGCGCTTAAGCGGAGCGCGGCTCTGCAAGCTTGCTTCC 257
QY 503 CCGAGTTGGCACCACCGAGAGGATGGGAGCCGACCCCTGACGTTCCAGAGGAGCCACCT 562
DB 256 CCGAGTTGGCACCACCGAGAGGATGGGAGCCGACCCCTGACGTTCCAGAGGAGCCACCT 197
QY 563 GAGAGCCAGGAGCGGTCAGAGACACGCTGTGACTCGAGATGCGGCTGCGGAGATGGA 622
DB 196 GAGAGCCAGGAGCGGTCAGAGACACGCTGTGACTCGAGATGCGGCTGCGGAGATGGA 137
QY 623 GAGGAGCGGAGGAGCGGCTTAACGGGCTCCCTCTGCGCGCCCGCTCCGAGAGGCGAC 682
DB 136 GAGGAGCGGAGGAGCGGCTTAACGGGCTCCCTCTGCGCGCCCGCTCCGAGAGGCGAC 77
QY 683 TCGAGGAGTCCCGGAGCGGCTCCGTGAGAGTGGCGGTAGCCGAGCGAGTCAAGAAC 742
DB 76 TCGAGGAGTCCCGGAGCGGCTCCGTGAGAGTGGCGGTAGCCGAGCGAGTCAAGAAC 17
QY 743 TGAAGAGCGTTCTGTC 758
DB 16 TGAAGAGCGTTCTGTC 1
```

## RESULT 14

```
US-09-925-065A-737120
/ Sequence 737120, Application US/09925065A
/ Publication No. US20050228172A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
```



```

; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 737120
; LENGTH: 555
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-737120

Query Match      17.8%; Score 554.6; DB 4; Length 555;
Best Local Similarity 99.8%; Pred. No. 1.1e-130;
Matches 554; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 204 CAGGCAAAATCTCCCGGAGAGCTGGCCGCTTTTAACTTGAGCTCAAGTTCCCATCC 263
DB 1 CAGGCAAAATCTCCCGGAGAGCTGGCCGCTTTTAACTTGAGCTCAAGTTCCCATCC 60

QY 264 GTAAATATGAACGAGGTGGATCTCCGAGCGCTTAACATTCAGAACTCGATGGGCGAA 323
DB 61 GTAAATATGAACGAGGTGGATCTCCGAGCGCTTAACATTCAGAACTCGATGGGCGAA 120

QY 324 GGGAGAGGAGGATGGGCGCAACCAAGTACCTCCCGCGTGGAGCCCGCTTACCACT 383
DB 121 GGGAGAGGAGGATGGGCGCAACCAAGTACCTCCCGCGTGGAGCCCGCTTACCACT 180

QY 384 GATTCAGAGGAGTGGAGCTCCGCGCGGAGCGAGCGGAGTGGGCGGCTTAAAGAAACCT 443
DB 181 GATTCAGAGGAGTGGAGCTCCGCGCGGAGCGAGCGGAGTGGGCGGCTTAAAGAAACCT 240

QY 444 ACCCGGCGCGCTTTGGCAGCGCTTAAGCGAGCGCGCTCTGAGCGCTTGGCCCG 503
DB 241 ACCCGGCGCGCTTTGGCAGCGCTTAAGCGAGCGCGCTCTGAGCGCTTGGCCCG 300

QY 504 GGAATTTGGACCAACCAAGATGGGAGCCGCACTTACGCTTGGAGGAGCGCAACCTG 563
DB 301 GGAATTTGGACCAACCAAGATGGGAGCCGCACTTACGCTTGGAGGAGCGCAACCTG 360

QY 564 AGGCGCAGGCGGAGTGGAGAGCAAGAGTGTGACTCGAGTGGCGCTCGGAGAGATGAGCG 623
DB 361 AGGCGCAGGCGGAGTGGAGAGCAAGAGTGTGACTCGAGTGGCGCTCGGAGAGATGAGCG 420

QY 624 AGGAGCGGAGAGCAAGCTTAACGAGGCTCTCTGCGCGCCCGCTCGCAGAGCGCAGCT 683
DB 421 AGGAGCGGAGAGCAAGCTTAACGAGGCTCTCTGCGCGCCCGCTCGCAGAGCGCAGCT 480

QY 684 CAGAGGCTCCCGGCGGAGCTCCGCTGAGAGTGGCGGTTAGGCGGAGCGAGTCAACAT 743
DB 481 CAGAGGCTCCCGGCGGAGCTCCGCTGAGAGTGGCGGTTAGGCGGAGCGAGTCAACAT 540

QY 744 GAAAGAGCTTCTGTC 758
DB 541 GAAAGAGCTTCTGTC 555

RESULT 15
US-10-029-386-6988/c
; Sequence 6988, Application US/10029386
; Publication No. US20030194704A1
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR C
; FILE REFERENCE: ABOMICA-X-2
; CURRENT APPLICATION NUMBER: US/10/029,386
; CURRENT FILING DATE: 2001-12-20

```

```

; NUMBER OF SEQ ID NOS: 34288
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO: 6988
; LENGTH: 524
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC010615.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 0.56
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3
; OTHER INFORMATION: EST_HUMAN HIT: AM302149.1, EVALU8 0.00e+00
; OTHER INFORMATION: SWISSPROT HIT: Q06805, EVALU8 1.90e-01
; OTHER INFORMATION: NT HIT: G14507086, EVALU8 7.80e+00
US-10-029-386-6988

Query Match      16.8%; Score 524; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 6.7e-123;
Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 484 CTTCGACCTGCTTGCCCGGAGTGGACCAACGAGATGGGAGCCGACCTTCAGC 543
DB 524 CTTCGACCTGCTTGCCCGGAGTGGACCAACGAGATGGGAGCCGACCTTCAGC 465

QY 544 TTCGAGGAGGACCAACCTGAGAGGCGGCGGTGAGAGCAACGAGTGTGATCTGAGT 603
DB 464 TTCGAGGAGGACCAACCTGAGAGGCGGCGGTGAGAGCAACGAGTGTGATCTGAGT 405

QY 604 GCGCTTGGGAGGATGAGCGAGGAGCGGAGGAGCCGCTTAAGGAGGCTCTGCGCGC 663
DB 404 GCGCTTGGGAGGATGAGCGAGGAGCGGAGGAGCCGCTTAAGGAGGCTCTGCGCGC 345

QY 664 CCGTCCGAGAGGCGCAAGTGAAGGCTCCGAGCGGCTCTGAGAGCTTGGCGGTACG 723
DB 344 CCGTCCGAGAGGCGCAAGTGAAGGCTCCGAGCGGCTCTGAGAGCTTGGCGGTACG 285

QY 724 CCGAGCGATCAACGAGCAATGAAGGCTTGTGTCGCGCGGCGCCCAAGCCGGATGAGG 783
DB 284 CCGAGCGATCAACGAGCAATGAAGGCTTGTGTCGCGCGGCGCCCAAGCCGGATGAGG 225

QY 784 GTTAGCCATCTCGCGCGCTGAGAGGAGGAGCTTAACGAGGCGGAGCGGCGCCAGC 843
DB 224 GTTAGCCATCTCGCGCGCTGAGAGGAGGAGCTTAACGAGGCGGAGCGGCGCCAGC 165

QY 844 CCGAGCCCAACCGGATGAGAGGAGGAGAGTCAAGGCGCTGAGAGCGGAGCTCAAGAG 903
DB 164 CCGAGCCCAACCGGATGAGAGGAGGAGAGTCAAGGCGCTGAGAGCGGAGCTCAAGAG 105

QY 904 GAGTGGCTGTCAACCAACCTGAGTGTGAGACCGTGTGAGCTCGCGGAGCTCGCAAGAC 963
DB 104 GAGTGGCTGTCAACCAACCTGAGTGTGAGACCGTGTGAGCTCGCGGAGCTCGCAAGAC 45

QY 964 TGCGGAGAGGCTGCAAAAGACGCGCGCAAGAGGCGCAGAGCTG 1007
DB 44 TGCGGAGAGGCTGCAAAAGACGCGCGCAAGAGGCGCAGAGCTG 1

Search completed: May 11, 2006, 03:56:35
Job time : 2410 secs

```

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using bw model

Run on: May 11, 2006, 02:14:07 ; Search time 1415 Seconds  
(without alignments)  
8998.684 Million cell updates/sec

Title: US-10-760-320A-102

Perfect score: 3122  
Sequence: 1 actagagcgggggttagcgcc.....acagagcaagactctgtctc 3122

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 9312410 seqs, 2039259788 residues

Total number of hits satisfying chosen parameters: 18624820

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications NA New:\*

- 1: /SIDSS/ptodata/1/pubpna/US06\_NEW\_PUB.seq1:\*
- 2: /SIDSS/ptodata/1/pubpna/US06\_NEW\_PUB.seq2:\*
- 3: /SIDSS/ptodata/1/pubpna/US07\_NEW\_PUB.seq1:\*
- 4: /SIDSS/ptodata/1/pubpna/US08\_NEW\_PUB.seq1:\*
- 5: /SIDSS/ptodata/1/pubpna/US08\_NEW\_PUB.seq2:\*
- 6: /SIDSS/ptodata/1/pubpna/US09\_NEW\_PUB.seq1:\*
- 7: /SIDSS/ptodata/1/pubpna/US09\_NEW\_PUB.seq2:\*
- 8: /SIDSS/ptodata/1/pubpna/US09\_NEW\_PUB.seq3:\*
- 9: /SIDSS/ptodata/1/pubpna/US10\_NEW\_PUB.seq1:\*
- 10: /SIDSS/ptodata/1/pubpna/US10\_NEW\_PUB.seq2:\*
- 11: /SIDSS/ptodata/1/pubpna/US10\_NEW\_PUB.seq3:\*
- 12: /SIDSS/ptodata/1/pubpna/US10\_NEW\_PUB.seq4:\*
- 13: /SIDSS/ptodata/1/pubpna/US10\_NEW\_PUB.seq5:\*
- 14: /SIDSS/ptodata/1/pubpna/US11\_NEW\_PUB.seq1:\*
- 15: /SIDSS/ptodata/1/pubpna/US11\_NEW\_PUB.seq2:\*
- 16: /SIDSS/ptodata/1/pubpna/US11\_NEW\_PUB.seq3:\*
- 17: /SIDSS/ptodata/1/pubpna/US11\_NEW\_PUB.seq4:\*
- 18: /SIDSS/ptodata/1/pubpna/US11\_NEW\_PUB.seq5:\*
- 19: /SIDSS/ptodata/1/pubpna/US60\_NEW\_PUB.seq1:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the total being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	559.6	17.9	560	7	US-09-925-065A-740956 Sequence 740956,
2	559.6	17.8	557	7	US-09-925-065A-769508 Sequence 769508,
3	555.6	17.8	556	7	US-09-925-065A-736351 Sequence 736351,
4	554.6	17.8	555	7	US-09-925-065A-737120 Sequence 737120,
5	180.2	5.8	2306	7	US-09-925-065A-706227 Sequence 706227,
6	178.6	5.7	414	7	US-09-925-065A-132933 Sequence 132933,
7	176.2	5.6	154548	17	US-11-121-086-33 Sequence 33, Appl
8	176	5.6	433	11	US-10-301-480-91009 Sequence 91009, A
9	176	5.6	433	11	US-10-301-480-91010 Sequence 91010, A
10	176	5.6	433	12	US-10-301-480-704418 Sequence 704418,
11	176	5.6	433	12	US-10-301-480-704419 Sequence 704419,
12	175.2	5.6	534	7	US-09-925-065A-64658 Sequence 64658, A
13	175.2	5.6	534	11	US-10-301-480-165896 Sequence 165896,
14	175.2	5.6	534	12	US-10-301-480-779305 Sequence 779305,

15	174.8	5.6	1541	7	US-09-925-065A-89171	Sequence 89171, A
16	174.8	5.6	1541	7	US-09-925-065A-89172	Sequence 89172, A
17	174.8	5.6	1541	11	US-10-301-480-190412	Sequence 190412,
18	174.8	5.6	1541	11	US-10-301-480-190413	Sequence 190413,
19	174.8	5.6	1541	12	US-10-301-480-803821	Sequence 803821,
20	174.8	5.6	1541	12	US-10-301-480-803822	Sequence 803822,
21	174.8	5.6	121736	18	US-11-114-798-49	Sequence 49, Appl
22	174.8	5.6	318488	18	US-11-114-798-58	Sequence 58, Appl
23	174.4	5.6	1600	7	US-09-925-065A-84740	Sequence 84740, A
24	174.4	5.6	1600	11	US-10-301-480-185960	Sequence 185960,
25	174.4	5.6	1600	11	US-10-301-480-799389	Sequence 799389,
26	174.2	5.6	1691140	17	US-11-091-018-1	Sequence 1, Appl1
27	174	5.6	516	7	US-09-925-065A-489267	Sequence 489267,
28	174	5.6	550	7	US-09-925-065A-546135	Sequence 546135,
29	174	5.6	550	12	US-10-301-480-524085	Sequence 524085,
30	174	5.6	550	12	US-10-301-480-1137494	Sequence 1137494,
31	174	5.6	150481	17	US-11-112-908-37	Sequence 37, Appl
32	174	5.6	155515	17	US-11-112-908-42	Sequence 42, Appl
33	174	5.6	159660	17	US-11-112-908-43	Sequence 43, Appl
34	174	5.6	171162	17	US-11-112-908-38	Sequence 38, Appl
35	174	5.6	177623	17	US-11-112-908-41	Sequence 41, Appl
36	173.6	5.6	386	12	US-10-301-480-229094	Sequence 229094,
37	173.6	5.6	386	12	US-10-301-480-842503	Sequence 842503,
38	173.6	5.6	717	7	US-09-925-065A-939937	Sequence 939937,
39	173.6	5.6	717	7	US-09-925-065A-955036	Sequence 955036,
40	173.6	5.6	908	12	US-10-301-480-609462	Sequence 609462,
41	173.6	5.6	908	12	US-10-301-480-1222871	Sequence 1222871,
42	173.4	5.6	1549	7	US-09-925-065A-726968	Sequence 726968,
43	173.4	5.6	101046	10	US-10-995-561-13330	Sequence 13330, A
44	173.2	5.5	679	12	US-10-301-480-531347	Sequence 531347,
45	173.2	5.5	679	12	US-10-301-480-1144756	Sequence 1144756,

#### ALIGNMENTS

RESULT 1  
US-09-925-065A-740956/C  
; Sequence 740956, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925, 065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243, 096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252, 147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250, 092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261, 766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289, 846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 740956  
; LENGTH: 560  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; US-09-925-065A-740956

Query Match 17.9%; Score 559.6; DB 7; Length 560;  
Best Local Similarity 99.8%; Pred. No. 0.017;  
Matches 559; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 199 CATCACAGCAAAACTCCGCGAGACCTGCGCCGTTTACCTGAGGCTCAGTTCC 258  
DB 560 CATCACAGCAAAACTCCGCGAGACCTGCGCCGTTTACCTGAGGCTCAGTTCC 501

OY	25	CATCCGTAATAATGAAACGGGTGGATCTCCGAGGCGTAACTTCAGAACTCCGAAATGG	318
Db	500	CATCCGTAATAATGAAACGGGTGGATCTCCGAGGCGTAACTTCAGAACTCCGAAATGG	441
OY	319	GCGAAGGGGAGGAGAGATGGGCAACCAACATGACTCCCGCGTGGAGCCCGGCTTA	378
Db	440	GCGAAGGGGAGGAGAGATGGGCAACCAACATGACTCCCGCGTGGAGCCCGGCTTA	381
OY	379	CCAATGATTCAGGGGGGTGGCAGTCCGGCCGGGAGAGATCGGGGTGGGGTCTTAAGAA	438
Db	380	CCAATGATTCAGGGGGGTGGCAGTCCGGCCGGGAGAGATCGGGGTGGGGTCTTAAGAA	322
OY	439	ACCTTAACCGGCGGCGCTTGGCAGCGCTTAAGGCGAGCGCGGCTTGCAGGCTGTT	498
Db	320	ACCTTAACCGGCGGCGCTTGGCAGCGCTTAAGGCGAGCGCGGCTTGCAGGCTGTT	261
OY	499	GCCCCGGGATTTGGCACCACAGAGATGGGGACCGCACCTCTGACTTTGGCAGGGAGCAC	558
Db	260	GCCCCGGGATTTGGCACCACAGAGATGGGGACCGCACCTCTGACTTTGGCAGGGAGCAC	201
OY	559	CGTGGAGGCGCAGGCGGCTGCAGAGACGACGCTGTGACTCGAGTGGCGCTGGGAGAGAT	618
Db	200	CGTGGAGGCGCAGGCGGCTGCAGAGACGACGCTGTGACTCGAGTGGCGCTGGGAGAGAT	141
OY	619	GGAGAGAGGAGCGGGGGGACCGCTTAACGGGGCTTCCTTGGCGCCCCGTCCGCAAGGCG	678
Db	140	GGAGAGAGGAGCGGGGGGACCGCTTAACGGGGCTTCCTTGGCGCCCCGTCCGCAAGGCG	81
OY	679	CACGTCCAGAGGATCCCGGAGCGGCTCCGTGGACGTTTGGCGGTCGACCGAGAGTCAAG	738
Db	80	CACGTCCAGAGGATCCCGGAGCGGCTCCGTGGACGTTTGGCGGTCGACCGAGAGTCAAG	21
OY	739	ACCATGAAGAGCGTTCTGTGC	758
Db	20	ACCATGAAGAGCGTTCTGTGC	1

```

RESULT 2
US-09-925-065A-769508/c
/ Sequence 769508, Application US/09925065A
/ Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
/ TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925, 065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243, 096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250, 092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261, 766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289, 846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 769508
LENGTH: 557
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-769508

```

Query Match	17.8%	Score 556.6	DB 7	Length 557
Best Local Similarity	99.8%	Pred. 0.002		
Matches 556, Conservative	1	Mismatches	0	Gaps 0
Qy	202	CACAGGCAAACTCCGCGGAGCCCTGGCCCGCTTTTAACTGGGAGCTACGTTTCCCAT	261	
Db	557	CACAGGCAAACTCCGCGGAGCCCTGGCCCGCTTTTAACTGGGAGCTACGTTTCCCAT	498	

QY 282 CCGTAAATATGAAACGGGTGGATCTCTCCCGAGGGCTTAATCAATTCACAGAACTCCGATATGGGGCG 321

Db 497 CCGTAAATATGAAACGGGTGGATCTCTCCCGAGGGCTTAATTCACAGAACTCCGATATGGGGCG 438

QY 322 AAGGGAGAGGAGGAGTGGGCGCACCCACACAGTGAAGCTCCCGCGTGGAGCCCGGCTTACCA 381

Db 437 AAGGGAGAGGAGGAGTGGGCGCACCCACACAGTGAAGCTCCCGCGTGGAGCCCGGCTTACCA 378

QY 382 CTGATCTCAGGGGGGTGGCAGCTTCCTCGGCGGGGAGAGCGGGGTGGCGGGGTCTTAGGAAATCC 441

Db 377 CTGATCTCAGGGGGGTGGCAGCTTCCTCGGCGGGGAGAGCGGGGTGGCGGGGTCTTAGGAAATCC 318

QY 442 CTACCCGGCCGGCTTTGGCAGCGCTTAAGCGGAGGCGCGGCTCTGACAGCTTGCTTGGC 501

Db 317 CTACCCGGCCGGCTTTGGCAGCGCTTAAGCGGAGGCGCGGCTCTGACAGCTTGCTTGGC 258

QY 502 CCGGAGTTGGCACCCACGAGAGATGGGAGCCGCACTTCAAGCTTCGACAGGAGACCAAGCT 561

Db 257 CCGGAGTTGGCACCCACGAGAGATGGGAGCCGCACTTCAAGCTTCGACAGGAGACCAAGCT 198

QY 562 GGAAGGCCAGAGCGGCTGCAGAGACACGACGCTGTGACTCGAGAGTGGCGCTTGGGAGAGATGGA 621

Db 197 GGAAGGCCAGAGCGGCTGCAGAGACACGACGCTGTGACTCGAGAGTGGCGCTTGGGAGAGATGGA 138

QY 622 CGAGGGAGCGGGGGAGACCGCTTAACGGGGGCTTCCTTCGCGGGCCCTTCCTCCCAAGAGGGCGAC 681

Db 137 CGAGGGAGCGGGGGAGACCGCTTAACGGGGGCTTCCTTCGCGGGCCCTTCCTCCCAAGAGGGCGAC 78

QY 682 GTCGAGGGTCCGGGCGGGGCTCCGTGACGTTTGCGGGTATGCGCGGACGAGATCACCGAAC 741

Db 77 GTCGAGGGTCCGGGCGGGGCTCCGTGACGTTTGCGGGTATGCGCGGACGAGATCACCGAAC 18

QY 742 ATGAAGAGCGTTCTGTC 758

Db 17 ATGAAGAGCGTTCTGTC 1

```

RESULT 3
US-09-925-065A-736351/c
Sequence 736351, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925, 065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243, 096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252, 147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250, 092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261, 766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289, 846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 736351
LENGTH: 556
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-736351

```

Query Match	17.8%	Score 555.6	DB 7	Length 556
Best Local Similarity	99.8%	Pred. No. 0.02		
Matches 555; Conservative	1	Mismatches 0	Indels 0	Gaps 0
QY	203	ACAGGCAAAACTCCGGGAGACCTTGGCCCTTTTACCTGGGCTCAGTTTCCCATC	262	

Db 556 ACAGCAAAAACCTCCGCGGAGCTGCCCCGCTTTTAACTTGCCCTCAGTTTCCCCATC 497  
Qy 263 CTTAAATAGAAAGGTTGATCTCCGAGCGCTAACATTCCAGAACTGGATGGGGCGA 322  
Db 496 CTTAAATAGAAAGGTTGATCTCCGAGCGCTTAATTCCAGAACTGGATGGGGCGA 437  
Qy 323 AGGGAGGAGGATGGGCAACCAACGTGACCTCCCGCGTGAAGCCCGCTAACAC 382  
Db 436 AGGGAGGAGGATGGGCAACCAACGTGACCTCCCGCGTGAAGCCCGCTAACAC 377  
Qy 383 TGAATCAGGGGTGGCACTCCGCGCGGAGCAAGCGGGGTGGCGGGTCTTAAAGAAACC 442  
Db 376 TGAATCAGGGGTGGCACTCCGCGCGGAGCAAGCGGGGTGGCGGGTCTTAAAGAAACC 317  
Qy 443 TACCCGCGCCCTTGGCAAGCGCTTAAGCGGAGCGCGGCTGTCAGACCTGCTTGCCC 502  
Db 316 TACCCGCGCCCTTGGCAAGCGCTTAAGCGGAGCGCGGCTGTCAGACCTGCTTGCCC 257  
Qy 503 CGAGTTGGCAACCAACGAGATGGGGAACCGACCTTCAAGCTTCGAGGAGCAACCGTG 562  
Db 256 CGAGTTGGCAACCAACGAGATGGGGAACCGACCTTCAAGCTTCGAGGAGCAACCGTG 197  
Qy 553 AGGCCAGGGCGGTGAGAGACACGAGTGTGACTCGAGTGGCGCTTGAGAGATGAC 622  
Db 196 AGGCCAGGGCGGTGAGAGACACGAGTGTGACTCGAGTGGCGCTTGAGAGATGAC 137  
Qy 623 GAGGAGCGGGGAGCCGCTTAAGCGGGCTCCCTTGCGCGCCCGCTCCGAGAGCGCACG 682  
Db 136 GAGGAGCGGGGAGCCGCTTAAGCGGGCTCCCTTGCGCGCCCGCTCCGAGAGCGCACG 77  
Qy 683 TCGAGGGTCCCGGCGGGCTCCGCTGAGCGTTGGCGGTGACGCGAGAGATGACCA 742  
Db 76 TCGAGGGTCCCGGCGGGCTCCGCTGAGCGTTGGCGGTGACGCGAGAGATGACCA 17  
Qy 743 TGAAGAGCGTTCGTGC 758  
Db 16 TGAAGAGCGTTCGTGC 1

RESULT 4  
US-09-925-065A-737120  
; Sequence 737120, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 737120  
; LENGTH: 555  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-737120

Query Match 17.8%; Score 554.6; DB 7; Length 555;  
Best Local Similarity 99.8%; Pred. No. 0.021;  
Matches 554; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
Qy 204 CAGCAAAAACCTCCGCGGAGCGCTGCTTTTAACTTGCCCTCAGTTTCCCCATC 263

Db 1 CAGCAAAAACCTCCGCGGAGCGCTGCCCCGCTTTTAACTTGCCCTCAGTTTCCCCATC 60  
Qy 264 CTTAAATAGAAAGGTTGATCTCCGAGCGCTTAATTCCAGAACTGGATGGGGCGA 323  
Db 61 CTTAAATAGAAAGGTTGATCTCCGAGCGCTTAATTCCAGAACTGGATGGGGCGA 120  
Qy 324 AGGGAGGAGGATGGGCAACCAACGTGACCTCCCGCGTGAAGCCCGCTAACAC 383  
Db 121 AGGGAGGAGGATGGGCAACCAACGTGACCTCCCGCGTGAAGCCCGCTAACAC 180  
Qy 384 GATTCAGGGGTGGCACTCCGCGCGGAGCAAGCGGGGTGGCGGGTCTTAAAGAAACC 443  
Db 181 GATTCAGGGGTGGCACTCCGCGCGGAGCAAGCGGGGTGGCGGGTCTTAAAGAAACC 240  
Qy 444 ACCCGCGCCCTTGGCAAGCGCTTAAGCGGAGCGCGGCTGTCAGACCTGCTTGCCC 503  
Db 241 ACCCGCGCCCTTGGCAAGCGCTTAAGCGGAGCGCGGCTGTCAGACCTGCTTGCCC 300  
Qy 504 GAGTTGGCAACCAACGAGATGGGGAACCGACCTTCAAGCTTCGAGGAGCAACCGTG 563  
Db 301 GAGTTGGCAACCAACGAGATGGGGAACCGACCTTCAAGCTTCGAGGAGCAACCGTG 360  
Qy 564 AGGCCAGGGCGGTGAGAGACACGAGTGTGACTCGAGTGGCGCTTGAGAGATGAC 623  
Db 361 AGGCCAGGGCGGTGAGAGACACGAGTGTGACTCGAGTGGCGCTTGAGAGATGAC 420  
Qy 624 AGGAGCGGGGAGCCGCTTAAGCGGGCTCCCTTGCGCGCCCGCTCCGAGAGCGCACG 683  
Db 421 AGGAGCGGGGAGCCGCTTAAGCGGGCTCCCTTGCGCGCCCGCTCCGAGAGCGCACG 480  
Qy 684 CGAGGGTCCCGGCGGGCTCCGCTGAGCGTTGGCGGTGACGCGAGAGATGACCA 743  
Db 481 CGAGGGTCCCGGCGGGCTCCGCTGAGCGTTGGCGGTGACGCGAGAGATGACCA 540  
Qy 744 GAAAGCGTTCGTGC 758  
Db 541 GAAAGCGTTCGTGC 555

RESULT 5  
US-09-925-065A-706227/c  
; Sequence 706227, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 706227  
; LENGTH: 2306  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-706227

Query Match 5.8%; Score 180.2; DB 7; Length 2306;  
Best Local Similarity 84.6%; Pred. No. 5.5e+03;  
Matches 214; Conservative 0; Mismatches 38; Indels 1; Gaps 1;



```
Db 307 CCCAGCACTTTGGAGGCTGAGGCGGAGAGATCACTGAGTCAAGAGTTTGAACACAGC 248
QY 2930 CTGGCCAAATATGCGAAACCCGATCTCTAATAAAATATAAAATTTGGCCGGGCAATGGT 2989
Db 247 CTGGCCAAATATGTAAGAACCCCACTCTCTTAAATAATACAAAATTAAGCTGGGCAATGGT 188
QY 2990 GCGCGATGCTGTGTGCCAGCTACTCGGAGAGTTGAGGCAAGAGATGCTTGAATGCA 3049
Db 187 GCGSACGCTGTAAATCCCTGTCTACTCAGAGAGCTGAGGCAAGAGATCACTTGAACCCA 128
QY 3050 GAGGTGAGGTTGCAATG-AGCAAGATTGTGCTGCACTGCACCTTGGGCAACAGAG 3108
Db 127 GAGGTGAGGTTGCAATGAGAGATGATGCTGCACTGCACCTTGGGCAACAGAG 68
QY 3109 CAAGACTCTGTCTC 3122
Db 67 CAAGACTCCATCTC 54

RESULT 9
US-10-301-480-91010/c
; Sequence 91010, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91010
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-91010

Query Match 5.6%; Score 176; DB 11; Length 433;
Best Local Similarity 83.1%; Pred. No. 2.3e+04;
Matches 211; Conservative 1; Mismatches 41; Indels 1; Gaps 1;

QY 2870 CCCCACTTTTCTTTTGGAGGCAAGTGATCACTGAGGCCAGAGTTTGAACACAGC 2929
Db 307 CCCAGCACTTTGGAGGCTGAGGCGGAGAGATCACTGAGTCAAGAGTTTGAACACAGC 248
QY 2930 CTGGCCAAATATGCGAAACCCGATCTCTAATAAAATATAAAATTTGGCCGGGCAATGGT 2989
Db 247 CTGGCCAAATATGTAAGAACCCCACTCTCTTAAATAATACAAAATTAAGCTGGGCAATGGT 188
QY 2990 GCGCGATGCTGTGTGCCAGCTACTCGGAGAGTTGAGGCAAGAGATGCTTGAATGCA 3049
Db 187 GCGSACGCTGTAAATCCCTGTCTACTCAGAGAGCTGAGGCAAGAGATCACTTGAACCCA 128
QY 3050 GAGGTGAGGTTGCAATG-AGCAAGATTGTGCTGCACTGCACCTTGGGCAACAGAG 3108
Db 127 GAGGTGAGGTTGCAATGAGAGATGATGCTGCACTGCACCTTGGGCAACAGAG 68
QY 3109 CAAGACTCTGTCTC 3122
Db 67 CAAGACTCCATCTC 54

RESULT 10
US-10-301-480-704418/c
; Sequence 704418, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
```

```
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE OF INVENTION: In the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 704418
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-704418

Query Match 5.6%; Score 176; DB 12; Length 433;
Best Local Similarity 83.1%; Pred. No. 2.3e+04;
Matches 211; Conservative 1; Mismatches 41; Indels 1; Gaps 1;

QY 2870 CCCCACTTTTCTTTTGGAGGCAAGTGATCACTGAGGCCAGAGTTTGAACACAGC 2929
Db 307 CCCAGCACTTTGGAGGCTGAGGCGGAGAGATCACTGAGTCAAGAGTTTGAACACAGC 248
QY 2930 CTGGCCAAATATGCGAAACCCGATCTCTAATAAAATATAAAATTTGGCCGGGCAATGGT 2989
Db 247 CTGGCCAAATATGTAAGAACCCCACTCTCTTAAATAATACAAAATTAAGCTGGGCAATGGT 188
QY 2990 GCGCGATGCTGTGTGCCAGCTACTCGGAGAGTTGAGGCAAGAGATGCTTGAATGCA 3049
Db 187 GCGSACGCTGTAAATCCCTGTCTACTCAGAGAGCTGAGGCAAGAGATCACTTGAACCCA 128
QY 3050 GAGGTGAGGTTGCAATG-AGCAAGATTGTGCTGCACTGCACCTTGGGCAACAGAG 3108
Db 127 GAGGTGAGGTTGCAATGAGAGATGATGCTGCACTGCACCTTGGGCAACAGAG 68
QY 3109 CAAGACTCTGTCTC 3122
Db 67 CAAGACTCCATCTC 54

RESULT 11
US-10-301-480-704419/c
; Sequence 704419, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 704419
; LENGTH: 433
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-704419

Query Match 5.6%; Score 176; DB 12; Length 433;
Best Local Similarity 83.1%; Pred. No. 2.3e+04;
Matches 211; Conservative 1; Mismatches 41; Indels 1; Gaps 1;

QY 2870 CCCCACTTTTCTTTTGGAGGCAAGTGATCACTGAGGCCAGAGTTTGAACACAGC 2929
Db 307 CCCAGCACTTTGGAGGCTGAGGCGGAGAGATCACTGAGTCAAGAGTTTGAACACAGC 248
```







**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2006, 03:56:47 ; Search time 15378 Seconds  
(without alignments)  
11540.204 Million cell updates/sec

Title: US-10-760-320A-102

Perfect score: 3122

Sequence: 1 actagaggtggtggttagcgc.....acagagcaagactctgctc 3122

Scoring table: OLIGO\_NWC

Gapop 60.0 , Gapext 60.0

Searched: 5683141 seqs, 28421725653 residues

Word size: 1

Total number of hits satisfying chosen parameters: 11765110

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: listing first 500 summaries

Database :

GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_in:\*  
3: gb\_env:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pr:\*  
9: gb\_pro:\*  
10: gb\_atc:\*  
11: gb\_gy:\*  
12: gb\_un:\*  
13: gb\_vl:\*  
14: gb\_hgt:\*  
15: gb\_pl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3122	100.0	3122	6	COR41455 Sequence
2	3122	100.0	3122	8	AK124439 Homo sapi
3	2840	91.0	191925	8	AC008736 Homo sapi
4	2840	91.0	216441	14	AC027340 Homo sapi
5	2445	78.3	186115	8	AC008474 Homo sapi
6	1952	62.5	2327	8	COS50499 Sequence
7	1952	62.5	2327	8	AK127646 Homo sapi
8	657	21.0	708	6	CQ748891 Homo sapi
9	646	20.7	655	8	HUM2D85E06 Homo sapi
10	606	19.4	708	6	CQ746021 Sequence
11	392	12.6	216441	14	AC027340 Homo sapi
12	144	4.6	449	10	G37408 SHGC-57769
13	60	1.9	60	6	COS53695 Sequence
14	53	1.7	24495	8	CR762434 Human DNA
15	53	1.7	42398	8	DB4401 Homo sapien
16	53	1.7	49878	8	CR547129 Human DNA
17	53	1.7	62449	14	AC061985 Homo sapi
18	53	1.7	69808	8	AC026771 Homo sapi

19	53	1.7	70621	8	AL713971 Human DNA
20	53	1.7	90476	8	AC015910 Homo sapi
21	53	1.7	100000	8	AP000208 Homo sapi
22	53	1.7	109523	8	AL844527 Homo sapi
23	53	1.7	110130	8	AL360268 Human DNA
24	53	1.7	113919	8	AP000247 Homo sapi
25	53	1.7	120825	8	HSJ733M16 Human DNA
26	53	1.7	134184	8	AC095044 Homo sapi
27	53	1.7	157686	8	BS000110 Pan trogl
28	53	1.7	156033	8	AL645940 Human DNA
29	53	1.7	171051	8	AC022960 Homo sapi
30	53	1.7	175737	8	HS1033B10 Human DNA
31	53	1.7	179462	14	AC021102 Homo sapi
32	53	1.7	185000	14	AC007332 Homo sapi
33	53	1.7	189181	14	AP001333 Homo sapi
34	53	1.7	190092	14	AC120502 Lemur cat
35	53	1.7	203583	8	AC092437 Homo sapi
36	53	1.7	204373	14	CR847805 Homo sapi
37	53	1.7	208220	8	AC009107 Homo sapi
38	53	1.7	218872	14	AC092434 Homo sapi
39	53	1.7	229090	14	CR847784 Homo sapi
40	53	1.7	323185	14	CR753824 Homo sapi
41	53	1.7	340000	8	AP001710 Homo sapi
42	53	1.7	349980	6	CS039425 Sequence
43	53	1.7	381	6	BD027540 Sequence
44	53	1.7	381	6	AX887930 Sequence
45	52	1.7	428	10	AB135735 Homo sapi
46	52	1.7	690	10	BV640976 Homo sapi
47	52	1.7	762	10	BV533342 Homo sapi
48	52	1.7	770	10	BV536183 Homo sapi
49	52	1.7	773	10	BV588891 G591P6387
50	52	1.7	775	10	BV492390 S221P6065
51	52	1.7	810	10	BV522612 G591P6007
52	52	1.7	915	10	BV465101 G591P6193
53	52	1.7	22738	8	AB014078 Homo sapi
54	52	1.7	23481	8	BX248096 Homo sapi
55	52	1.7	25057	8	AC087649 Homo sapi
56	52	1.7	25242	8	AL356788 Human DNA
57	52	1.7	25972	8	CR336916 Human DNA
58	52	1.7	28276	8	CR382280 Human DNA
59	52	1.7	32906	8	CR391992 Human DNA
60	52	1.7	38173	8	AC093235 Homo sapi
61	52	1.7	39372	8	HSN121B8 Human DNA s
62	52	1.7	39372	8	HSN121B8 Human DNA s
63	52	1.7	39596	14	AL445257 Homo sapi
64	52	1.7	41372	8	AL359082 Human DNA
65	52	1.7	42572	8	AC004602 Homo sapi
66	52	1.7	44449	8	AC016627 Homo sapi
67	52	1.7	44783	8	CR368927 Homo sapi
68	52	1.7	46201	8	AC006139 Homo sapi
69	52	1.7	46509	8	AC108747 Homo sapi
70	52	1.7	47462	8	CR759912 Homo sapi
71	52	1.7	51090	14	AL390837 Homo sapi
72	52	1.7	59744	14	AC103989 Homo sapi
73	52	1.7	62761	8	AL499606 Homo sapi
74	52	1.7	64056	8	AC092080 Homo sapi
75	52	1.7	66887	8	AP411057 Homo sapi
76	52	1.7	68959	8	HSJ735P11 Homo sapi
77	52	1.7	70393	8	AC073488 Homo sapi
78	52	1.7	70448	8	CR388372 Human DNA
79	52	1.7	73597	14	AC109598 Homo sapi
80	52	1.7	75525	8	AL391561 Human DNA
81	52	1.7	80117	8	AC055813 Homo sapi
82	52	1.7	82724	8	BX927220 Homo sapi
83	52	1.7	92139	6	AR242855 Sequence
84	52	1.7	92139	6	AX384907 Sequence
85	52	1.7	92442	14	CR751233 Homo sapi
86	52	1.7	93790	8	AP397423 Homo sapi
87	52	1.7	96217	8	AL645729 Human DNA
88	52	1.7	97630	8	AL844213 Human DNA
89	52	1.7	101769	14	AC093203 Homo sapi
90	52	1.7	102311	8	AC005377 Homo sapi
91	52	1.7	105902	8	AL359699 Human DNA

92	52	1.7	105956	8	AC074136	155	52	1.7	186415	8	AC008675	AC008675 Homo sapi
93	52	1.7	106539	8	AP225899	156	52	1.7	186418	8	AC018757	AC018757 Homo sapi
94	52	1.7	110000	8	BA000025_14	167	52	1.7	186448	14	AC034129	AC034129 Homo sapi
95	52	1.7	110000	8	BA000041_09	168	52	1.7	188596	8	AC120045	AC120045 Homo sapi
96	52	1.7	110000	8	BA000041_10	169	52	1.7	188640	14	AC146153	AC146153 Pan trogl
97	52	1.7	110028	8	HS1164110	170	52	1.7	191037	8	AC091561	AC091561 Homo sapi
98	52	1.7	110042	8	AL590788	171	52	1.7	191234	14	AC018681	AC018681 Homo sapi
99	52	1.7	111051	14	AF322449	172	52	1.7	191540	8	AC144780	AC144780 Pan trogl
100	52	1.7	116888	8	AC114481	173	52	1.7	191866	14	AC068676	AC068676 Homo sapi
101	52	1.7	117899	8	AC016590	174	52	1.7	191924	14	AC091440	AC091440 Homo sapi
102	52	1.7	120723	8	AC104051	175	52	1.7	192328	8	AL954212	AL954212 Pan trogl
103	52	1.7	128540	14	AC025473	176	52	1.7	192505	8	AL353136	AL353136 Homo sapi
104	52	1.7	130146	8	AL590489	177	52	1.7	193351	14	AC110601	AC110601 Homo sapi
105	52	1.7	130572	14	AC024096	178	52	1.7	193952	14	AC141263	AC141263 Homo sapi
106	52	1.7	130984	14	AC022766	179	52	1.7	194189	8	AC103705	AC103705 Homo sapi
107	52	1.7	134465	8	AC100793	180	52	1.7	195388	8	AP002026	AP002026 Homo sapi
108	52	1.7	135805	14	AC069511	181	52	1.7	196622	8	AC125238	AC125238 Homo sapi
109	52	1.7	136385	8	AC006006	182	52	1.7	196623	8	AC099058	AC099058 Homo sapi
110	52	1.7	137845	14	AC010367	183	52	1.7	198105	14	AC026160	AC026160 Homo sapi
111	52	1.7	138573	8	AL954211	184	52	1.7	200000	14	AC008108	AC008108 Homo sapi
112	52	1.7	141633	8	AC073542	185	52	1.7	200853	8	CNS01DSQ	AL121839 Human chr
113	52	1.7	143369	8	AL357552	186	52	1.7	201167	8	AL162591	AL162591 Human DNA
114	52	1.7	144542	14	AC027146	187	52	1.7	202539	8	AC090559	AC090559 Homo sapi
115	52	1.7	146199	14	AC110281	188	52	1.7	202971	14	AC151042	AC151042 Callithr
116	52	1.7	146376	8	AC009247	189	52	1.7	204917	8	AC008040	AC008040 Homo sapi
117	52	1.7	148049	8	CNS01DPTX	190	52	1.7	205283	8	AC073916	AC073916 Homo sapi
118	52	1.7	148049	8	HSN300188	191	52	1.7	205736	14	AC023560	AC023560 Homo sapi
119	52	1.7	149679	8	AC002351	192	52	1.7	207661	8	AC015884	AC015884 Homo sapi
120	52	1.7	149776	8	AC012070	193	52	1.7	211812	14	AC027205	AC027205 Homo sapi
121	52	1.7	150159	8	AC018791	194	52	1.7	213947	14	AC135988	AC135988 Homo sapi
122	52	1.7	150437	6	CS086345	195	52	1.7	215780	14	AC135989	AC135989 Homo sapi
123	52	1.7	150846	8	AL138724	196	52	1.7	234053	8	AC002429	AC002429 Homo sapi
124	52	1.7	151228	8	AL662800	197	52	1.7	256000	14	AC145895	AC145895 Pan trogl
125	52	1.7	151970	14	AC040905	198	52	1.7	279430	14	AC145065	AC145065 Pan trogl
126	52	1.7	152040	14	AC139547	199	52	1.7	349980	6	CS039416	CS039416 Sequence
127	52	1.7	152492	8	AC079863	200	52	1.6	715	10	BV634964	BV634964 Human DNA
128	52	1.7	157135	14	AC064817	201	51	1.6	14814	8	AL357035	AL357035 Human DNA
129	52	1.7	160929	8	AC008677	202	51	1.6	34869	14	CR318623	CR318623 Homo sapi
130	52	1.7	161870	14	AC027706	203	51	1.6	68003	14	BX323853	BX323853 Homo sapi
131	52	1.7	162948	8	AL662822	204	51	1.6	71173	8	AC069148	AC069148 Homo sapi
132	52	1.7	163157	8	AC108670	205	51	1.6	86882	8	BX649553	BX649553 Human DNA
133	52	1.7	163380	14	AC141308	206	51	1.6	93663	8	AC022409	AC022409 Homo sapi
134	52	1.7	163432	14	AC161476	207	52	1.6	105392	8	AC073215	AC073215 Homo sapi
135	52	1.7	163562	14	AC092606	208	51	1.6	125403	8	AC008929	AC008929 Homo sapi
136	52	1.7	163562	14	AC096508	209	51	1.6	126682	8	AC008649	AC008649 Homo sapi
137	52	1.7	164429	14	AC130837	210	51	1.6	146193	8	AC027347	AC027347 Homo sapi
138	52	1.7	165311	8	AC097061	211	51	1.6	148782	8	AP006213	AP006213 Homo sapi
139	52	1.7	165972	14	AC135854	212	51	1.6	153998	14	AC013329	AC013329 Homo sapi
140	52	1.7	166484	14	AC026830	213	51	1.6	154625	14	CR936360	CR936360 Homo sapi
141	52	1.7	167780	8	AL844892	214	51	1.6	154754	8	AC064836	AC064836 Homo sapi
142	52	1.7	167920	8	AC109635	215	51	1.6	159392	14	AC148829	AC148829 Pan trogl
143	52	1.7	168734	8	AC007371	216	51	1.6	163246	14	AC068401	AC068401 Homo sapi
144	52	1.7	168793	8	AC025787	217	51	1.6	165433	8	AC146088	AC146088 Pan trogl
145	52	1.7	169059	14	AC009790	218	51	1.6	169627	8	AL136984	AL136984 Human DNA
146	52	1.7	169893	14	AC021165	219	51	1.6	174034	8	AC020908	AC020908 Homo sapi
147	52	1.7	170232	8	AL135903	220	51	1.6	174428	8	AC005821	AC005821 Homo sapi
148	52	1.7	171058	8	AC073068	221	51	1.6	175501	8	AC008427	AC008427 Homo sapi
149	52	1.7	171849	8	AC010616	222	51	1.6	175994	8	AC079460	AC079460 Homo sapi
150	52	1.7	172307	8	AL590609	223	51	1.6	177777	14	AC068862	AC068862 Homo sapi
151	52	1.7	172759	8	AC025589	224	51	1.6	178665	14	AP003483	AP003483 Homo sapi
152	52	1.7	172915	14	AC010160	225	51	1.6	180133	8	AL138885	AL138885 Human DNA
153	52	1.7	174034	8	AC020908	226	51	1.6	184133	8	AC103559	AC103559 Homo sapi
154	52	1.7	174521	14	AC127468	227	51	1.6	184289	8	AL359878	AL359878 Human DNA
155	52	1.7	176343	8	CNS01DX3	228	51	1.6	184569	8	AC010583	AC010583 Homo sapi
156	52	1.7	176871	14	AC150824	229	51	1.6	186239	8	AC123982	AC123982 Homo sapi
157	52	1.7	177744	14	AC073954	230	51	1.6	190708	8	AC005261	AC005261 Homo sapi
158	52	1.7	177773	8	AC010761	231	51	1.6	191311	14	AC091486	AC091486 Homo sapi
159	52	1.7	178168	14	AC012350	232	51	1.6	192826	8	AC113266	AC113266 Homo sapi
160	52	1.7	181792	8	AC093126	233	51	1.6	194296	8	AL354864	AL354864 Human DNA
161	52	1.7	182152	14	AC067929	234	51	1.6	195766	8	CNS01DVC	AL135744 Human chr
162	52	1.7	182892	8	AC034244	235	51	1.6	198927	14	AC026373	AC026373 Homo sapi
163	52	1.7	183556	14	AC019059	236	51	1.6	219122	8	AC080038	AC080038 Homo sapi
164	52	1.7	184252	14	AC139565	237	51	1.6	227137	8	AC098481	AC098481 Homo sapi

C 238	50	1.6	255	6	BD048209	311	50	1.6	125465	8	HSJ136J15	AL118496 Human DNA
C 239	50	1.6	255	6	AX912676	312	50	1.6	128450	8	AC108861	AC108861 Homo sapi
C 240	50	1.6	686	10	BV664432	C 313	50	1.6	129472	8	AL451125	AL451125 Human DNA
C 241	50	1.6	726	10	BV617582	C 314	50	1.6	129600	14	AL163539	AL163539 Homo sapi
C 242	50	1.6	730	10	BV622197	C 315	50	1.6	130436	8	HS17K7	AL350188 Human DNA
C 243	50	1.6	817	10	BV668130	C 316	50	1.6	130554	14	AL391537	AL391537 Homo sapi
C 244	50	1.6	911	10	BV632403	C 317	50	1.6	130647	8	AL157385	AL157385 Human DNA
C 245	50	1.6	1001	6	CO853998	C 318	50	1.6	130706	14	AC024123	AC024123 Homo sapi
C 246	50	1.6	1258	8	BC030749	C 319	50	1.6	130843	8	HS1798J3	AL035603 Human DNA
C 247	50	1.6	2188	8	AK025471	C 320	50	1.6	139111	8	AL158048	AL158048 Human DNA
C 248	50	1.6	3597	8	HUMAP011A	C 321	50	1.6	139846	8	AC020895	AC020895 Homo sapi
C 249	50	1.6	3625	8	AK130932	C 322	50	1.6	140091	8	AL445684	AL445684 Human DNA
C 250	50	1.6	8705	6	ARS67482	C 323	50	1.6	140195	8	AC087163	AC087163 Homo sapi
C 251	50	1.6	8705	6	ARS69796	C 324	50	1.6	140455	8	AC010719	AC010719 Homo sapi
C 252	50	1.6	8705	6	AX277531	C 325	50	1.6	141670	8	AC010417	AC010417 Homo sapi
C 253	50	1.6	8705	6	AX418095	C 326	50	1.6	141899	8	AC087457	AC087457 Homo sapi
C 254	50	1.6	15169	8	AC134915	C 327	50	1.6	142225	8	AC110053	AC110053 Homo sapi
C 255	50	1.6	23536	8	AC002120	C 328	50	1.6	145340	8	AL136525	AL136525 Human DNA
C 256	50	1.6	35865	8	AC011552	C 329	50	1.6	145770	14	AC019244	AC019244 Homo sapi
C 257	50	1.6	40338	14	AC147414	C 330	50	1.6	146466	8	AL353637	AL353637 Human DNA
C 258	50	1.6	48723	8	AL645992	C 331	50	1.6	149103	8	AC013434	AC013434 Homo sapi
C 259	50	1.6	51770	8	AL590430	C 332	50	1.6	149559	8	AL139327	AL139327 Human DNA
C 260	50	1.6	52368	8	AC010621	C 333	50	1.6	150228	8	AC003071	AC003071 Homo sapi
C 261	50	1.6	55672	14	AC103991	C 334	50	1.6	150708	8	AP003467	AP003467 Homo sapi
C 262	50	1.6	62164	8	AL670729	C 335	50	1.6	151479	14	AC012278	AC012278 Homo sapi
C 263	50	1.6	63437	8	HSJ964417	C 336	50	1.6	151660	14	AC027450	AC027450 Homo sapi
C 264	50	1.6	64475	8	AP28332062	C 337	50	1.6	152210	8	AY445095	AY445095 Homo sapi
C 265	50	1.6	64989	14	AC024899	C 338	50	1.6	152846	14	AC025709	AC025709 Homo sapi
C 266	50	1.6	65268	14	AC099847	C 339	50	1.6	152905	14	AC044801	AC044801 Homo sapi
C 267	50	1.6	65958	8	AC105413	C 340	50	1.6	154567	14	AL591375	AL591375 Human DNA
C 268	50	1.6	69926	8	AL593851	C 341	50	1.6	154592	8	AL591215	AL591215 Human DNA
C 269	50	1.6	71577	8	AC135587	C 342	50	1.6	155067	8	AC010301	AC010301 Homo sapi
C 270	50	1.6	72402	8	AL389924	C 343	50	1.6	156094	8	AC090198	AC090198 Homo sapi
C 271	50	1.6	73656	14	AL016512	C 344	50	1.6	157207	8	BX640542	BX640542 Human DNA
C 272	50	1.6	77001	8	AL513013	C 345	50	1.6	157828	8	BX005266	BX005266 Human DNA
C 273	50	1.6	79666	8	AL136969	C 346	50	1.6	157921	14	AC007850	AC007850 Homo sapi
C 274	50	1.6	80427	8	AC010492	C 347	50	1.6	158255	14	AC027726	AC027726 Homo sapi
C 275	50	1.6	81490	8	AL590987	C 348	50	1.6	158825	14	AC025695	AC025695 Homo sapi
C 276	50	1.6	82901	8	AL391335	C 349	50	1.6	158902	14	AC069481	AC069481 Homo sapi
C 277	50	1.6	84031	8	AP000223	C 350	50	1.6	158929	14	AC105089	AC105089 Homo sapi
C 278	50	1.6	86800	8	AC055715	C 351	50	1.6	159098	14	AC021980	AC021980 Homo sapi
C 279	50	1.6	88698	14	AC009008	C 352	50	1.6	160193	8	AC092442	AC092442 Homo sapi
C 280	50	1.6	90399	14	AL162405	C 353	50	1.6	161264	8	AC007011	AC007011 Homo sapi
C 281	50	1.6	92487	8	AC068280	C 354	50	1.6	161306	14	AC091056	AC091056 Homo sapi
C 282	50	1.6	92872	8	AC022425	C 355	50	1.6	161507	8	AC046195	AC046195 Homo sapi
C 283	50	1.6	93381	14	HSJ0766D4	C 356	50	1.6	161791	14	AC021690	AC021690 Homo sapi
C 284	50	1.6	95097	8	AC022202	C 357	50	1.6	162100	8	AC018868	AC018868 Homo sapi
C 285	50	1.6	96299	8	AL133519	C 358	50	1.6	163167	8	AC011347	AC011347 Homo sapi
C 286	50	1.6	96489	8	AC011395	C 359	50	1.6	163520	8	AL391259	AL391259 Human DNA
C 287	50	1.6	97687	8	AL591403	C 360	50	1.6	163589	14	AC008470	AC008470 Homo sapi
C 288	50	1.6	98078	8	AL451083	C 361	50	1.6	163599	8	HSJ081D8	HSJ081D8 Human DNA
C 289	50	1.6	98842	8	AC008972	C 362	50	1.6	163954	8	AC090959	AC090959 Homo sapi
C 290	50	1.6	99505	8	AP002087	C 363	50	1.6	164118	14	AC021384	AC021384 Homo sapi
C 291	50	1.6	100000	8	AP000085	C 364	50	1.6	164236	8	BS000105	BS000105 Pan trogl
C 292	50	1.6	100000	8	AP000137	C 365	50	1.6	164529	14	AC079975	AC079975 Homo sapi
C 293	50	1.6	100000	8	AC114962	C 366	50	1.6	164766	14	AC026657	AC026657 Homo sapi
C 294	50	1.6	107714	14	AL356306	C 367	50	1.6	164774	14	AC019104	AC019104 Homo sapi
C 295	50	1.6	108162	14	AC020960	C 368	50	1.6	164933	14	AC074231	AC074231 Homo sapi
C 296	50	1.6	110000	14	AC008576	C 369	50	1.6	165773	8	AL350737	AL350737 Human DNA
C 297	50	1.6	110000	14	AC079517	C 370	50	1.6	165942	8	AC104943	AC104943 Homo sapi
C 298	50	1.6	110000	14	AL590046_2	C 371	50	1.6	166687	8	BX284677	BX284677 Human DNA
C 299	50	1.6	110753	8	AC010216	C 372	50	1.6	166876	14	AL356864	AL356864 Homo sapi
C 300	50	1.6	112707	8	AL359207	C 373	50	1.6	166895	8	AC147030	AC147030 Pan trogl
C 301	50	1.6	112954	14	AP302687	C 374	50	1.6	167013	8	HSJ088H9	HSJ088H9 Human DNA
C 302	50	1.6	114148	8	AL312655	C 375	50	1.6	167170	8	AC120782	AC120782 Pan trogl
C 303	50	1.6	114191	8	AL451007	C 376	50	1.6	167293	14	AC022629	AC022629 Homo sapi
C 304	50	1.6	114540	8	AC107426	C 377	50	1.6	167386	8	AL159993	AL159993 Homo sapi
C 305	50	1.6	114983	14	AC093309	C 378	50	1.6	168111	8	HSJ251L6	HSJ251L6 Human DNA
C 306	50	1.6	117520	8	AP005624	C 379	50	1.6	168114	8	AL354828	AL354828 Homo sapi
C 307	50	1.6	120129	8	AL391560	C 380	50	1.6	168435	8	AC104429	AC104429 Homo sapi
C 308	50	1.6	121282	8	AC027318	C 381	50	1.6	168493	8	AC013463	AC013463 Homo sapi
C 309	50	1.6	123369	8	AC025170	C 382	50	1.6	168499	8	AC092037	AC092037 Homo sapi
C 310	50	1.6	124102	8	AC021148	C 383	50	1.6	168672	14	AC027102	AC027102 Homo sapi



384	50	1.6 168753	8	AC006120	AC006120 Homo sapi	457	50	1.6 196590	14	AC156839	AC156839 Pan trogl
385	50	1.6 168873	14	BXS37112	BXS37112 Homo sapi	458	50	1.6 197337	6	AR658697	AR658697 Sequence
386	50	1.6 168875	8	AC114794	AC114794 Homo sapi	459	50	1.6 197819	8	AC152891	AC152891 Pan trogl
387	50	1.6 168880	8	AC0117104	AC0117104 Homo sapi	460	50	1.6 198017	8	AC012314	AC012314 Homo sapi
388	50	1.6 169166	8	AC027687	AC027687 Homo sapi	461	50	1.6 198027	14	AC008515	AC008515 Homo sapi
389	50	1.6 169218	14	AC149013	AC149013 Papio anu	462	50	1.6 198253	8	AC009090	AC009090 Homo sapi
390	50	1.6 169835	8	AP005717	AP005717 Homo sapi	463	50	1.6 198700	8	AC104317	AC104317 Homo sapi
391	50	1.6 170349	8	AC068799	AC068799 Homo sapi	464	50	1.6 198720	8	AP002365	AP002365 Homo sapi
392	50	1.6 171951	8	CNS051DA	ALJ36020 Human chr	465	50	1.6 198899	8	AC090192	AC090192 Homo sapi
393	50	1.6 172288	8	AC046181	AC046181 Homo sapi	466	50	1.6 199218	8	AC108863	AC108863 Homo sapi
394	50	1.6 172286	14	AL160164	AL160164 Homo sapi	467	50	1.6 199566	8	AC103676	AC103676 Homo sapi
395	50	1.6 172514	14	AC146903	AC146903 Aotus nan	468	50	1.6 200000	14	AC004624	AC004624 Homo sapi
396	50	1.6 172753	8	ALJ35929	ALJ35929 Human DNA	469	50	1.6 200000	14	AC004670	AC004670 Homo sapi
397	50	1.6 172816	8	AC093889	AC093889 Homo sapi	470	50	1.6 201414	8	AC090644	AC090644 Homo sapi
398	50	1.6 172837	8	ALJ450304	ALJ450304 Human DNA	471	50	1.6 202891	8	EX088651	EX088651 Human DNA
399	50	1.6 173333	14	AC025282	AC025282 Homo sapi	472	50	1.6 203838	14	AC026840	AC026840 Homo sapi
400	50	1.6 173706	14	AC022694	AC022694 Homo sapi	473	50	1.6 204102	14	AC013826	AC013826 Homo sapi
401	50	1.6 173952	14	AC080087	AC080087 Homo sapi	474	50	1.6 204439	8	AP000807	AP000807 Homo sapi
402	50	1.6 174023	8	AC080089	AC080089 Homo sapi	475	50	1.6 204505	8	AC089983	AC089983 Homo sapi
403	50	1.6 174097	8	AC069513	AC069513 Homo sapi	476	50	1.6 205307	14	AC009968	AC009968 Homo sapi
404	50	1.6 174612	14	AC023889	AC023889 Homo sapi	477	50	1.6 206258	14	AC112775	AC112775 Homo sapi
405	50	1.6 174693	8	ALJ36738	ALJ36738 Human DNA	478	50	1.6 207815	14	AC053497	AC053497 Homo sapi
406	50	1.6 174995	8	ALJ365502	ALJ365502 Human DNA	479	50	1.6 208117	8	AC008488	AC008488 Homo sapi
407	50	1.6 175021	14	AC161086	AC161086 Papio ham	480	50	1.6 209532	14	ALJ36136	ALJ36136 Homo sapi
408	50	1.6 175347	14	AC009099	AC009099 Homo sapi	481	50	1.6 209888	14	AC146462	AC146462 Salimiri b
409	50	1.6 175440	14	AC137526	AC137526 Homo sapi	482	50	1.6 210643	8	AC137723	AC137723 Homo sapi
410	50	1.6 175493	14	AC020769	AC020769 Homo sapi	483	50	1.6 210643	8	AC137723	AC137723 Homo sapi
411	50	1.6 175822	14	AC068402	AC068402 Homo sapi	484	50	1.6 211211	14	AC161277	AC161277 Pan trogl
412	50	1.6 176550	8	ALJ363536	ALJ363536 Human DNA	485	50	1.6 211213	14	AC163792	AC163792 Homo sapi
413	50	1.6 176647	8	AC024163	AC024163 Homo sapi	486	50	1.6 211550	8	ALJ62430	ALJ62430 Homo sapi
414	50	1.6 176671	14	AC080149	AC080149 Homo sapi	487	50	1.6 211791	8	ALJ62430	ALJ62430 Homo sapi
415	50	1.6 176708	14	AC027548	AC027548 Homo sapi	488	50	1.6 212622	8	AC025259	AC025259 Homo sapi
416	50	1.6 176932	8	AC016772	AC016772 Homo sapi	489	50	1.6 212978	8	AC007383	AC007383 Homo sapi
417	50	1.6 177448	8	AC044781	AC044781 Homo sapi	490	50	1.6 216129	14	AC157530	AC157530 Pan trogl
418	50	1.6 177518	8	AC093577	AC093577 Homo sapi	491	50	1.6 216129	8	AC023825	AC023825 Homo sapi
419	50	1.6 177555	8	AC006450	AC006450 Homo sapi	492	50	1.6 221887	8	ALJ58052	ALJ58052 Human DNA
420	50	1.6 177688	8	CNS01DR5	ALJ32640 Human chr	493	50	1.6 226910	14	AC141314	AC141314 Homo sapi
421	50	1.6 177846	14	AC010893	AC010893 Homo sapi	494	50	1.6 233418	14	AC109128	AC109128 Homo sapi
422	50	1.6 178384	14	AC011885	AC011885 Homo sapi	495	50	1.6 234287	6	AR659646	AR659646 Sequence
423	50	1.6 179176	14	AC022857	AC022857 Homo sapi	496	50	1.6 234288	6	AR659600	AR659600 Sequence
424	50	1.6 181123	14	AC149133	AC149133 Pan trogl	497	50	1.6 241178	14	AC120839	AC120839 Pan trogl
425	50	1.6 181237	8	AC104064	AC104064 Homo sapi	498	50	1.6 243722	14	AC151855	AC151855 Papio ham
426	50	1.6 181598	8	AC026172	AC026172 Homo sapi	499	50	1.6 250529	8	HUAE00658	HUAE00658 Homo sapi
427	50	1.6 181604	8	AC026194	AC026194 Homo sapi	500	50	1.6 295639	14	ALJ60019	ALJ60019 Homo sapi
428	50	1.6 181663	8	AC010894	AC010894 Homo sapi	ALIGNMENTS					
429	50	1.6 182567	14	AC145207	AC145207 Homo sapi	RESULT 1					
430	50	1.6 182945	14	AC069004	AC069004 Homo sapi	LOCUS					
431	50	1.6 182996	8	CNS01DRU	ALJ32986 Human chr	DEFINITION					
432	50	1.6 183353	8	AC087071	AC087071 Homo sapi	AC084155					
433	50	1.6 183539	8	AC092268	AC092268 Homo sapi	AC084155					
434	50	1.6 183818	14	AC149455	AC149455 Papio anu	AC084155					
435	50	1.6 183842	8	AC130450	AC130450 Homo sapi	AC084155					
436	50	1.6 184059	14	AC024196	AC024196 Homo sapi	AC084155					
437	50	1.6 184536	14	AC068682	AC068682 Homo sapi	AC084155					
438	50	1.6 184706	8	AC120781	AC120781 Pan trogl	AC084155					
439	50	1.6 184759	8	AC022695	AC022695 Homo sapi	AC084155					
440	50	1.6 184906	8	CNS01DRQ	ALJ33313 Human chr	AC084155					
441	50	1.6 186615	8	EX088717	EX088717 Human DNA	AC084155					
442	50	1.6 187126	14	AP002395	AP002395 Homo sapi	AC084155					
443	50	1.6 187239	14	AC021349	AC021349 Homo sapi	AC084155					
444	50	1.6 187343	14	AC016688	AC016688 Homo sapi	AC084155					
445	50	1.6 188636	8	HUAC00287	HUAC00287 Homo sapi	AC084155					
446	50	1.6 190368	8	AC146080	AC146080 Pan trogl	AC084155					
447	50	1.6 191668	14	CR759969	CR759969 Homo sapi	AC084155					
448	50	1.6 192478	14	AC164657	AC164657 Pan trogl	AC084155					
449	50	1.6 193882	8	AFJ33041	AFJ33041 Homo sapi	AC084155					
450	50	1.6 193889	14	AC151014	AC151014 Callithrix	AC084155					
451	50	1.6 194919	14	AC127899	AC127899 Homo sapi	AC084155					
452	50	1.6 195280	8	AC073869	AC073869 Homo sapi	AC084155					
453	50	1.6 195383	14	AC126472	AC126472 Homo sapi	AC084155					
454	50	1.6 196096	14	ALJ61451	ALJ61451 Homo sapi	AC084155					
455	50	1.6 196207	14	AC027048	AC027048 Homo sapi	AC084155					
456	50	1.6 196416	8	AC006483	AC006483 Homo sapi	AC084155					

Query Match	100.0%	Score 3122	DB 6	Length 3122
Best Local Similarity	100.0%	Pred. No. 0		
Matches 3122	Conservative 0	Mismatches 0	Indels 0	Gaps 0
QY	1	ACTAGAGGTGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGCAGAGAC	60	
Db	1	ACTAGAGGTGGGTTAGCGCTTGGAAGCACCGACCAACGTGAGCGCAACGCGCAGAGAC	60	
QY	61	ACCTGACCCCGCGCGCGCCAGCGCCCTCGGAATTGCACTGCTCGCTTTGGGGCAG	120	
Db	61	ACCTGACCCCGCGCGCGCCAGCGCCCTCGGAATTGCACTGCTCGCTTTGGGGCAG	120	
QY	121	GAGGTGCGCAGTCTGCGCGGGGACCCGAGGTCTGTGCGCGGACAGGGTCCGGAGTCACT	180	
Db	121	GAGGTGCGCAGTCTGCGCGGGGACCCGAGGTCTGTGCGCGGAGTCCGGAGTCACT	180	
QY	181	ATAGCTGGGTTCTAGTCCCATGACAGGCAAAAATCTCCGCGGAGGCTTGACCCGCTTTTAA	240	
Db	181	ATAGCTGGGTTCTAGTCCCATGACAGGCAAAAATCTCCGCGGAGGCTTGACCCGCTTTTAA	240	
QY	241	CTTGCGGCTCACTTTCCCATCCGTAATAGAACGGGTTGATCTCCCGAGCGCTAAACA	300	
Db	241	CTTGCGGCTCACTTTCCCATCCGTAATAGAACGGGTTGATCTCCCGAGCGCTAAACA	300	
QY	301	TTCCAGAACTCGGATGGGGCGAAAGGGGAGGAGGAGTGGGCAACCCAGACGTGACCTCC	360	
Db	301	TTCCAGAACTCGGATGGGGCGAAAGGGGAGGAGGAGTGGGCAACCCAGACGTGACCTCC	360	
QY	361	CGCGTGAAGCCCGCGCTACCACTGATTCAGGGGGTGGCAGCTCCGCGCGGACAGACGG	420	
Db	361	CGCGTGAAGCCCGCGCTACCACTGATTCAGGGGGTGGCAGCTCCGCGCGGACAGACGG	420	
QY	421	GTGGCGGGTCTTAGGAAACCTTACC CGGCGCCCTTGGCAGCGCTTAAAGCGAGCGG	480	
Db	421	GTGGCGGGTCTTAGGAAACCTTACC CGGCGCCCTTGGCAGCGCTTAAAGCGAGCGG	480	
QY	481	CGGCTCTGAGCGTGTGGCCCCGAGTTGGGACCCAGAGAGATGGGGACCGCACCCCTC	540	
Db	481	CGGCTCTGAGCGTGTGGCCCCGAGTTGGGACCCAGAGAGATGGGGACCGCACCCCTC	540	
QY	541	AGCTTCGACGAGGCCACCGTGAAGGCGAGGCGGTGACAGACCAAGATGTGACTCGG	600	
Db	541	AGCTTCGACGAGGCCACCGTGAAGGCGAGGCGGTGACAGACCAAGATGTGACTCGG	600	
QY	601	AGTGGCGCTGGGAGAGATGACGAGGAGCGGGGACCGCTTAAACGGGCTCCCTTGCGC	660	
Db	601	AGTGGCGCTGGGAGAGATGACGAGGAGCGGGGACCGCTTAAACGGGCTCCCTTGCGC	660	
QY	661	GGCGGTCGCGAGAGCGGACGTCGAGGGGTCCCGGCGGGCTCCGTCGACGTGGCGGTA	720	
Db	661	GGCGGTCGCGAGAGCGGACGTCGAGGGGTCCCGGCGGGCTCCGTCGACGTGGCGGTA	720	
QY	721	GCGCGAGCGAGTCAACGACCAATGAAGAGCTTTCGTGCGCGCGGCCCAAGGCGCGGATG	780	
Db	721	GCGCGAGCGAGTCAACGACCAATGAAGAGCTTTCGTGCGCGCGGCCCAAGGCGCGGATG	780	
QY	781	GGGGTTAGCCATCTCTGCGCGCTGAGGGGAGGGCTTAAACGGGCGCGGCGCGCGCC	840	
Db	781	GGGGTTAGCCATCTCTGCGCGCTGAGGGGAGGGCTTAAACGGGCGCGGCGCGCGCC	840	
QY	841	AGCGGAGGCGCAACCGGATGGGCGAGGAGAGATGCAAGGCGCTGTCGACCGGCTCAACA	900	
Db	841	AGCGGAGGCGCAACCGGATGGGCGAGGAGAGATGCAAGGCGCTGTCGACCGGCTCAACA	900	
QY	901	AGACGACTGCGTCTACCAACCACTTGCTGACCGTCTGCTGCTCGCGGACCTCGCAGA	960	
Db	901	AGACGACTGCGTCTACCAACCACTTGCTGACCGTCTGCTGCTCGCGGACCTCGCAGA	960	
QY	961	ACCTGGGCGAGAGCTGCAAAAAGACGCGCAAGAGGCGCAGAGGCTGGGGGTCTCAACT	1020	
Db	961	ACCTGGGCGAGAGCTGCAAAAAGACGCGCAAGAGGCGCAGAGGCTGGGGGTCTCAACT	1020	

QY	1021	GCGCCGCGTGAAGTCTGTGCTGCGGACCGGGGCGTGGCGGCGACGACGCGCGAGT	1080	
Db	1021	GCGCCGCGTGAAGTCTGTGCTGCGGACCGGGGCGTGGCGGCGACGACGCGCGAGT	1080	
QY	1081	TCGAGCGGCTCTGGGTGCGCTTCTCGGGCTGTCGACCTGTGGAAGCGGACATGCGAC	1140	
Db	1081	TCGAGCGGCTCTGGGTGCGCTTCTCGGGCTGTCGACCTGTGGAAGCGGACATGCGAC	1140	
QY	1141	GCTGCTGGAAGTGGGCGCGCGCTTCCGCTGCAAGCGCCCGGGGACCGCTGGTGGCA	1200	
Db	1141	GCTGCTGGAAGTGGGCGCGCGCTTCCGCTGCAAGCGCCCGGGGACCGCTGGTGGCA	1200	
QY	1201	CAGGTGGCTGAGCGCTCTTCCGCGGTGAGCGGCGCGCGCTGACACCCGACCTTGC	1260	
Db	1201	CAGGTGGCTGAGCGCTCTTCCGCGGTGAGCGGCGCGCGCTGACACCCGACCTTGC	1260	
QY	1261	GAGTGAAGCGGAGGCGCACTTCAAGTTCGCGGACCTTGCGGAGCTGAGCGCGAGTTC	1320	
Db	1261	GAGTGAAGCGGAGGCGCACTTCAAGTTCGCGGACCTTGCGGAGCTGAGCGCGAGTTC	1320	
QY	1321	TTCAAGTGGCGGAGTGAATCCACCAATGAGAGATGAAGTCAACGTGCCCGCTGACCG	1380	
Db	1321	TTCAAGTGGCGGAGTGAATCCACCAATGAGAGATGAAGTCAACGTGCCCGCTGACCG	1380	
QY	1381	TGCAAGCCCGGCGAGCGCGCGGCGCGAGCTCTGTCAACGCTCAAGCGCGCGCTCTCT	1440	
Db	1381	TGCAAGCCCGGCGAGCGCGCGGCGCGAGCTCTGTCAACGCTCAAGCGCGCGCTCTCT	1440	
QY	1441	CGGTGTGTCTTTGAGAGAGCGCGGGGAGGTTGCGACCCGAGAAAGCTTGCGCGCA	1500	
Db	1441	CGGTGTGTCTTTGAGAGAGCGCGGGGAGGTTGCGACCCGAGAAAGCTTGCGCGCA	1500	
QY	1501	TCCTTTTGGGCGCGGTGCTGCTGGGCGCTGGGCGCTTACCGGTGGGCGGAGCTGA	1560	
Db	1501	TCCTTTTGGGCGCGGTGCTGCTGGGCGCTGGGCGCTTACCGGTGGGCGGAGCTGA	1560	
QY	1561	GCTGACAGACACCCGACGCGCTGCTGCTGCGCTCCCTCCCTGAGAAAGAATCGG	1620	
Db	1561	GCTGACAGACACCCGACGCGCTGCTGCTGCGCTCCCTCCCTGAGAAAGAATCGG	1620	
QY	1621	GATGGGTGGGCTTGCGCTGTGCAAGGAGATGATCTTAAACCCGCTGTGTGATG	1680	
Db	1621	GATGGGTGGGCTTGCGCTGTGCAAGGAGATGATCTTAAACCCGCTGTGTGATG	1680	
QY	1681	GTAACAAGCGGCTTCCATGTCACATTCGCTGGGAGAGACACGCTTTCCTTTCGTGCG	1740	
Db	1681	GTAACAAGCGGCTTCCATGTCACATTCGCTGGGAGAGACACGCTTTCCTTTCGTGCG	1740	
QY	1741	CCGGAGAAAGTTAATTTGCGCGCGCGCTGACAGGCAATTAACGCTTACAGTCAAGACT	1800	
Db	1741	CCGGAGAAAGTTAATTTGCGCGCGCGCTGACAGGCAATTAACGCTTACAGTCAAGACT	1800	
QY	1801	TTATTCCTTAATTAAGAAAACCGTCAAGTGAACCTTGAATCCCTCCAGTTAATGAGTT	1860	
Db	1801	TTATTCCTTAATTAAGAAAACCGTCAAGTGAACCTTGAATCCCTCCAGTTAATGAGTT	1860	
QY	1861	AAACAATGCTGTTGGGCGCTTTCACAGGAGTCCGAGTTCCGCTGTCACCCGCGCA	1920	
Db	1861	AAACAATGCTGTTGGGCGCTTTCACAGGAGTCCGAGTTCCGCTGTCACCCGCGCA	1920	
QY	1921	GCGTGGCCCTTTCTGCGTGGGACAGTTTGAAGTGGGAGTGGAGTGAAGTTTG	1980	
Db	1921	GCGTGGCCCTTTCTGCGTGGGACAGTTTGAAGTGGGAGTGGAGTGAAGTTTG	1980	
QY	1981	GAGAGGAGCGCTGTTGTTCTAGTGGTGGTCTGTTTCCCGACCAAGAAAATTGCA	2040	
Db	1981	GAGAGGAGCGCTGTTGTTCTAGTGGTGGTCTGTTTCCCGACCAAGAAAATTGCA	2040	
QY	2041	TCAATGTCAAGAGCTTTTAATTAACCTTAATCTTCAAGGCGCTTAAATTTAGAGAGTGC	2100	
Db	2041	TCAATGTCAAGAGCTTTTAATTAACCTTAATCTTCAAGGCGCTTAAATTTAGAGAGTGC	2100	
QY	2101	TGAGAGAGTTCTATCAAAAGGCTTCTCTAAGACGCGCTACAGCCCTTCTTACAGAGT	2160	

```

Db      2101 TGAAGGAGCTTATCAAAAGGCGTTCTCTTAAGACGCGCTAGACGCTTCTCAGCAGAGT 2160
Qy      2161 TTATTCATTGTCGCCAAGAGAGAGAGATTTGAGGTCAATGCTCCCACTGCCG 2220
Db      2161 TTATTCATTGTCGCCAAGAGAGAGATTTGAGGTCAATGCTCCCACTGCCG 2220
Qy      2221 CTCAGGGGCTGACCCATTTAGAGAAACAAAGAGGGTGGTGAAGCTTCTCCAGCAG 2280
Db      2221 CTCAGGGGCTGACCCATTTAGAGAAACAAAGAGGGTGGTGAAGCTTCTCCAGCAG 2280
Qy      2281 TTGATTCAGTCGACGACCTTCCTGCGAGAAAGGGCTCTCCCAAGCACCAGAGATG 2340
Db      2281 TTGATTCAGTCGACGACCTTCCTGCGAGAAAGGGCTCTCCCAAGCACCAGAGATG 2340
Qy      2341 GGGTAAAGAGAGAGAGAGAGGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400
Db      2341 GGGTAAAGAGAGAGAGAGAGGCTTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400
Qy      2401 TCCCTCTCGGGGCTTATTTTGTTCAGAACTAGACAGAGGTTTGAACCTCCTTGA 2460
Db      2401 TCCCTCTCGGGGCTTATTTTGTTCAGAACTAGACAGAGGTTTGAACCTCCTTGA 2460
Qy      2461 GGAGGAGCTGGAGATCCTCTTTAGAGACCTTAACTCTTATTCCTCGAGATGTCGTG 2520
Db      2461 GGAGGAGCTGGAGATCCTCTTTAGAGACCTTAACTCTTATTCCTCGAGATGTCGTG 2520
Qy      2521 TGGCCAGTGAAGAGGCTGGCTTTGGAGCTCCCTGACCCCGCGCTGCCCGCTCCG 2580
Db      2521 TGGCCAGTGAAGAGGCTGGCTTTGGAGCTCCCTGACCCCGCGCTGCCCGCTCCG 2580
Qy      2581 GGTAAATGTGGCATTACTGGCCCAAGAGGTTTGAAGCAATAGCTGAGCTGGGTTA 2640
Db      2581 GGTAAATGTGGCATTACTGGCCCAAGAGGTTTGAAGCAATAGCTGAGCTGGGTTA 2640
Qy      2641 GAATGTAAAGCTTTAACTTTGAGATTTAAAGATTTAAAGATTTAAAGATTTAAAG 2700
Db      2641 GAATGTAAAGCTTTAACTTTGAGATTTAAAGATTTAAAGATTTAAAGATTTAAAG 2700
Qy      2701 GAAAAATGACGTAAACAGAGCTGTAATGAAGCTGTTATTTATTAAGAGAGCTGG 2760
Db      2701 GAAAAATGACGTAAACAGAGCTGTAATGAAGCTGTTATTTATTAAGAGAGCTGG 2760
Qy      2761 GCCATGAATCTATACCTGCAATGAGTCAACATGATCTTATGATGATCTTAGATT 2820
Db      2761 GCCATGAATCTATACCTGCAATGAGTCAACATGATCTTATGATGATCTTAGATT 2820
Qy      2821 ACTAATATATATTTATCTACTTCTGAGATTTGATGATTTCCCGCCCGCCCACTTTT 2880
Db      2821 ACTAATATATATTTATCTACTTCTGAGATTTGATGATTTCCCGCCCGCCCACTTTT 2880
Qy      2881 TCTTTTGGAGGAGAGTGAATGACCTGAGAGCCAGAGATTTGAGAGCAAGCTGGCCAA 2940
Db      2881 TCTTTTGGAGGAGAGTGAATGACCTGAGAGCCAGAGATTTGAGAGCAAGCTGGCCAA 2940
Qy      2941 AGCGAAACCCGATCTCTACTTAAATAATTAATAATTTGGCCGAGGAGTGGGAGTGC 3000
Db      2941 AGCGAAACCCGATCTCTACTTAAATAATTAATAATTTGGCCGAGGAGTGGGAGTGC 3000
Qy      3001 GTGGTCCAGCTACTCGGAGGTTGAGGAGGAGAGTGCCTTGAATGAGAGGTTGAG 3060
Db      3001 GTGGTCCAGCTACTCGGAGGTTGAGGAGGAGAGTGCCTTGAATGAGAGGTTGAG 3060
Qy      3061 TTGCAATGAGAGAGATTTGGCCATGCACTCAAGCTGGGCAACAGAGAGAGAGAG 3120
Db      3061 TTGCAATGAGAGAGATTTGGCCATGCACTCAAGCTGGGCAACAGAGAGAGAGAG 3120
Qy      3121 TC 3122
Db      3121 TC 3122

```

RESULT 2

```

AK124499
LOCUS      AK124499      3122 bp      mRNA      linear      PRI 09-SEP-2003
DEFINITION Homo sapiens CDNA FLJ42508 f18, clone BRACE204365.
ACCESSION AK124499
VERSION   AK124499.1 GI:34530295
KEYWORDS  oligo capping; f18 (full insert sequence).
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE
  1 Kawakami,B., Sugiyama,A., Takemoto,M., Sugiyama,T., Irie,R.,
    Otsuki,T., Sato,H., Wakamatsu,A., Iehi,S., Yamamoto,J., Isono,Y.,
    Kawai-Hio,Y., Saito,K., Nishikawa,T., Kimura,K., Yamashita,H.,
    Matsuo,K., Nakamura,Y., Sekine,M., Kikuchi,H., Kanda,K.,
    Magatsuma,M., Murakawa,K., Kasehori,K., Takahashi-Fuji,A.,
    Oshima,A., Suzuki,Y., Sugano,S., Nagahara,K., Masuno,Y., Nagai,K.
    and Isoqal,T.
    NEDO human cDNA sequencing project
    Unpublished
    2 (bases 1 to 3122)
    Isoqal,T. and Yamamoto,J.
    Direct Submission
    Submitted (15-JUN-2003) Takao Isoqal, Fuj Project (HRI Team) ; 2-6-7
    Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
    (E-mail:genomcse@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)
    NEDO human cDNA sequencing project supported by Ministry of
    Economy, Trade and Industry of Japan; cDNA full insert sequencing:
    Research Association for Biotechnology (RAB) / cDNA library
    construction: Helix Research Institute (HRI) (supported by Japan
    Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
    HRI, and Biotechnology Center, National Institute of Technology and
    Evaluation; clone selection for full insert sequencing: HRI and
    RAB; annotation: HRI and RAB.
FEATURES
  source
    1..3122
    location/Qualifiers
      1..3122
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="BRACE204365"
        /issue_type="cerebellum"
        /clone_id="BRACE2"
        /note="Cloning vector: pMB18F13"
        858..1565
        /note="unnamed protein product"
        /codon_start=1
        /protein_id="BAC85865.1"
        /db_xref="GI:34530296"
        /translation="MARBECKALADGINKTACVHLLVLTVGSGADSONLROELOKTR
          OKABIAVSTCARLTAVIDRGLAADRAFEERLWVAFSGCLDLIDMRSLRGLAA
          PPLAAPPRLVITGVAAAGSSVYARALSTRRLREABEDPRVADLRERVLVDGEM
          IDNEMKVNVRMTVQARQAAAGAEILSTVSGPSSVSLQERGGGCDPRKLAAILFG
          AVLLAAVALAVCAVAKLS"
ORIGIN
  Query Match      100.0%; Score 3122; DB 8; Length 3122;
  Best Local Similarity 100.0%; Pred. No. 0;
  Matches 3122; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  1 ACTAGAGGTGGGTTAGCGCTTGAAGAGCAGGACCAAGTGAAGGCAACGCGGAGGAGC 60
  Db 1 ACTAGAGGTGGGTTAGCGCTTGAAGAGCAGGACCAAGTGAAGGCAACGCGGAGGAGC 60
  Qy 61 ACTGACCCCGGCGGCGCCAGCCCTCGAATGGCACTGCTGCTTTGGGAGCAG 120
  Db 61 ACTGACCCCGGCGGCGCCAGCCCTCGAATGGCACTGCTGCTTTGGGAGCAG 120
  Qy 121 GAGGCGCCAGTCCCTGCGGGGCAACCCAGCGTCGTCGCGGCAAGGAGTCCGGAGTCA 180
  Db 121 GAGGCGCCAGTCCCTGCGGGGCAACCCAGCGTCGTCGCGGCAAGGAGTCCGGAGTCA 180
  Qy 181 ATAGCTGGGTTCTAGTCCCATCAGAGCAAAAGTCCGCGGAGGCTGGCCCGCTTTTAA 240

```

Db 181 ATAGCTGGGTTCTAAGTCCATCAAGCAAAAATCTCCGGGGAGCTTGGCCGCTTTTA 240  
Qy 241 CCTGGGCTCAAGTTTCCCATCCGTTAAATGAACGGGTTGATCTTCCGAGGCTTAACA 300  
Db 241 CCTGGGCTCAAGTTTCCCATCCGTTAAATGAACGGGTTGATCTTCCGAGGCTTAACA 300  
Qy 301 TTCCAGAACTCGGATGGGGGAAAGGGGAGGAGATGGGCAACCCACAGTGAACCTCC 360  
Db 301 TTCCAGAACTCGGATGGGGGAAAGGGGAGGAGATGGGCAACCCACAGTGAACCTCC 360  
Qy 361 CGGTTGAGAGCCCGCCTTACCTGATCAAGGGGTGGAGCTTCGCGCGGAGACAGCGAG 420  
Db 361 CGGTTGAGAGCCCGCCTTACCTGATCAAGGGGTGGAGCTTCGCGCGGAGACAGCGAG 420  
Qy 421 GTGGGGGGGTTCTTAAGAAACCTTACCCGCGCCCTTGGAGGCTTAAAGCGAGCGG 480  
Db 421 GTGGGGGGGTTCTTAAGAAACCTTACCCGCGCCCTTGGAGGCTTAAAGCGAGCGG 480  
Qy 481 CGGCTGTGAGCGCTGTGCGCCCGAGTTGGCAACCCAGAGGAGTGGGAGCGCACCTTC 540  
Db 481 CGGCTGTGAGCGCTGTGCGCCCGAGTTGGCAACCCAGAGGAGTGGGAGCGCACCTTC 540  
Qy 541 AGCTTCGAGGAGCCACCGTGGAGGCGAGGGCGGTGAGAGACAGAGCTGTGACTCGG 600  
Db 541 AGCTTCGAGGAGCCACCGTGGAGGCGAGGGCGGTGAGAGACAGAGCTGTGACTCGG 600  
Qy 601 AGTGGCCCTGGGGAGGATGAGCAGAGGAGCGGGAGACCGCTTAAACGGGGCTCCTTCGCGC 660  
Db 601 AGTGGCCCTGGGGAGGATGAGCAGAGGAGCGGGAGACCGCTTAAACGGGGCTCCTTCGCGC 660  
Qy 661 GGGCCGTCGAGAGGCGACGTCGAGGGGTCGCGGGCGGGCTCCGTTGGAGCGTTGGGGGTA 720  
Db 661 GGGCCGTCGAGAGGCGACGTCGAGGGGTCGCGGGCGGGCTCCGTTGGAGCGTTGGGGGTA 720  
Qy 721 GGGCCGAGGAGTCAAGGAGCCATGAAGAGGCTGTGCGCGCGCGCCAGAGCGGGATG 780  
Db 721 GGGCCGAGGAGTCAAGGAGCCATGAAGAGGCTGTGCGCGCGCGCCAGAGCGGGATG 780  
Qy 781 GGGGTTTAAACCATCTCTGCGCGCTGAGGGGAGGCTTAAACGGGCGGGCGCGCGCC 840  
Db 781 GGGGTTTAAACCATCTCTGCGCGCTGAGGGGAGGCTTAAACGGGCGGGCGCGCGCC 840  
Qy 841 AGCGGAGGCGCACCGGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900  
Db 841 AGCGGAGGCGCACCGGATGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 900  
Qy 901 AGACGACTCGTGTCTCAACCACTGTGTGACCGTGGGTGCTCGGCGAGACTCGCAGA 960  
Db 901 AGACGACTCGTGTCTCAACCACTGTGTGACCGTGGGTGCTCGGCGAGACTCGCAGA 960  
Qy 961 ACCCTGGGAGAGCTGCAGAAAGACGCGCGAGAGGCGGAGAGTGGCGGTGTCAACT 1020  
Db 961 ACCCTGGGAGAGCTGCAGAAAGACGCGCGAGAGGCGGAGAGTGGCGGTGTCAACT 1020  
Qy 1021 GGGCGCGGCTGACTGTGTGCTGGGAGACCGGGGCTTGGCGCGCGAGCGGCGGAGT 1080  
Db 1021 GGGCGCGGCTGACTGTGTGCTGGGAGACCGGGGCTTGGCGCGCGAGCGGCGGAGT 1080  
Qy 1081 TCGAGCGGCTTGGGTGAGCTTCTCGGGCTGCTGAGACCTGTGAAGGCGAGCATGCGAC 1140  
Db 1081 TCGAGCGGCTTGGGTGAGCTTCTCGGGCTGCTGAGACCTGTGAAGGCGAGCATGCGAC 1140  
Qy 1141 GCTCGCTGAGAGCTGGGCGCGCGCTTCCGCTGCAACGCGCGCGGAGCATGCTGCGCA 1200  
Db 1141 GCTCGCTGAGAGCTGGGCGCGCGCTTCCGCTGCAACGCGCGCGGAGCATGCTGCGCA 1200  
Qy 1201 CAGGTGTGCTGGCGCTCTCTCGGAGTGGCGGCGGCGGCGGCTGAGACCCCGAGCTTC 1260  
Db 1201 CAGGTGTGCTGGCGCTCTCTCGGAGTGGCGGCGGCGGCGGCTGAGACCCCGAGCTTC 1260  
Qy 1261 GGGTCGAGGCGAGGCGAGCTTTCGAGCTGCGGAGCTTCGCGAGAGCTGAGAGCGGAGTTC 1320

Db 1261 GGGTCGAGGCGAGGCGAGCTTTCGAGCTGCGGAGCTTCGCGAGAGCTGAGAGCGGAGTTC 1320  
Qy 1321 TTCAAGTGGGAGATGATGCAACAATGAGAGATGAAGGTCAACGTCGCCGCTGAGACG 1380  
Db 1321 TTCAAGTGGGAGATGATGCAACAATGAGAGATGAAGGTCAACGTCGCCGCTGAGACG 1380  
Qy 1381 TGCAGAGCCCGGAGGCGGCGGCGCGGAGCTCTGTGCAAGGTGACGCGGCGCCCTCT 1440  
Db 1381 TGCAGAGCCCGGAGGCGGCGGCGCGGAGCTCTGTGCAAGGTGACGCGGCGCCCTCT 1440  
Qy 1441 CGGTGCTGCTTTCGAGAGCGCGGGGGGGTTCGAGCCCAAGAAAGGCTTCGCGCGCA 1500  
Db 1441 CGGTGCTGCTTTCGAGAGCGCGGGGGGGTTCGAGCCCAAGAAAGGCTTCGCGCGCA 1500  
Qy 1501 CGGTGCTGCTTTCGAGAGCGCGGGGGGGTTCGAGCCCAAGAAAGGCTTCGCGCGCA 1500  
Db 1501 CGGTGCTGCTTTCGAGAGCGCGGGGGGGTTCGAGCCCAAGAAAGGCTTCGCGCGCA 1500  
Qy 1501 TCCCTTTTCGCGCGCGGCTGTGGGCGGTGAGCCCTTACCGGTGTGCGGAGACTGA 1560  
Db 1501 TCCCTTTTCGCGCGCGGCTGTGGGCGGTGAGCCCTTACCGGTGTGCGGAGACTGA 1560  
Qy 1561 GCTGACAGACACCGACCGGCGCTGTGCTGCTCCGCTCCCTGAGAAAGACTCGG 1620  
Db 1561 GCTGACAGACACCGACCGGCGCTGTGCTGCTCCGCTCCCTGAGAAAGACTCGG 1620  
Qy 1621 GATGGGTGTGGGTTCTGGCTGTGCAAGGAGAGTGTCTTAAACCCGCTGTGTGATG 1680  
Db 1621 GATGGGTGTGGGTTCTGGCTGTGCAAGGAGAGTGTCTTAAACCCGCTGTGTGATG 1680  
Qy 1681 GTTACAGCGGCTTTCAGTGTGCAATCTGCTTGGGAGAGCAAGGTTTCTTGTGTGCG 1740  
Db 1681 GTTACAGCGGCTTTCAGTGTGCAATCTGCTTGGGAGAGCAAGGTTTCTTGTGTGCG 1740  
Qy 1741 CCGGAGAAAGTTTAACTTTGCGCGCGGCGCTGAGGAGATTACCGCTTAAAGTCTGAGAGCT 1800  
Db 1741 CCGGAGAAAGTTTAACTTTGCGCGCGGCGCTGAGGAGATTACCGCTTAAAGTCTGAGAGCT 1800  
Qy 1801 TTATTCCTTAAATTAAGAAACCGTCAAGTACCTTAAAGTCTTAAAGTCTTAAAGT 1860  
Db 1801 TTATTCCTTAAATTAAGAAACCGTCAAGTACCTTAAAGTCTTAAAGTCTTAAAGT 1860  
Qy 1861 AACATGTGCTGTGGGCGGCTTTCAGAGGAGTCCGAGTTCGAGTCCGAGCGGCGCA 1920  
Db 1861 AACATGTGCTGTGGGCGGCTTTCAGAGGAGTCCGAGTTCGAGTCCGAGCGGCGCA 1920  
Qy 1921 GGGTGGCGGCTTTCGCGTGGAGCAGTTTGAAGGAGTGGGAGTGAAGTTCG 1980  
Db 1921 GGGTGGCGGCTTTCGCGTGGAGCAGTTTGAAGGAGTGGGAGTGAAGTTCG 1980  
Qy 1981 GAGAGGAGCGCTGTGTGCTTATGTGTGTGCTTCCGAGCAAGAAAGTTGCA 2040  
Db 1981 GAGAGGAGCGCTGTGTGCTTATGTGTGTGCTTCCGAGCAAGAAAGTTGCA 2040  
Qy 2041 TCAATGTGAGAGCTTTCATTAATCTTTCAGAGGCTTAAATTTAGAGAGTTC 2100  
Db 2041 TCAATGTGAGAGCTTTCATTAATCTTTCAGAGGCTTAAATTTAGAGAGTTC 2100  
Qy 2101 TGAAGCAGTTTCAAAAGGCTTTCCTAAGAGCGGCTTCAAGGCTTCTAGCAGAGT 2160  
Db 2101 TGAAGCAGTTTCAAAAGGCTTTCCTAAGAGCGGCTTCAAGGCTTCTAGCAGAGT 2160  
Qy 2161 TTATTCATTTCGTCGCAAGAGAGTGAAGAGTTCGAGTTCGAGTTCGAGTTCG 2220  
Db 2161 TTATTCATTTCGTCGCAAGAGAGTGAAGAGTTCGAGTTCGAGTTCGAGTTCG 2220  
Qy 2221 CTTCAGGGGCTGACCTTATTTAGAGAAACCAAGAGGAGTTCGAGTTCGAGTTCG 2280  
Db 2221 CTTCAGGGGCTGACCTTATTTAGAGAAACCAAGAGGAGTTCGAGTTCGAGTTCG 2280  
Qy 2281 TTGATTCAGTTCGAGCACTTCGCTGCGGAGAAAGGCTTCGCGAGCAACCGAGAGTTC 2340  
Db 2281 TTGATTCAGTTCGAGCACTTCGCTGCGGAGAAAGGCTTCGCGAGCAACCGAGAGTTC 2340  
Qy 2341 GGGTAAAGAGAGAGAGGCTTGGGGTGAAGGCGCACTGTGTTCGAGGCACTTTC 2400  
Db 2341 GGGTAAAGAGAGAGAGGCTTGGGGTGAAGGCGCACTGTGTTCGAGGCACTTTC 2400

QY	2401	TCCTCTCTGGGGGCTTATTTTGTTCAGAACTGACCCAGAGTGTGTAACCTCTTGCA	2460
Db	2401	TCCTCTCTGGGGGCTTATTTTGTTCAGAACTGACCCAGAGTGTGTAACCTCTTGCA	2460
QY	2461	GGAGGGCTGGGAATCTCTTTAGAGCACTTAATCTTATATCCCTGGAAATGTGCGTC	2520
Db	2461	GGAGGGCTGGGAATCTCTTTAGAGCACTTAATCTTATATCCCTGGAAATGTGCGTC	2520
QY	2521	TGGCCAGTAGAGAGGGCTGGCTTTTGGAGCTCCCTGACCCCGCGCTGCCCTCCCG	2580
Db	2521	TGGCCAGTAGAGAGGGCTGGCTTTTGGAGCTCCCTGACCCCGCGCTGCCCTCCCG	2580
QY	2581	GCTAATGTGGGATTACTGGCCCAAGAGTTTGTAGCAATCAGCTCTGAGACTGGGTTA	2640
Db	2581	GCTAATGTGGGATTACTGGCCCAAGAGTTTGTAGCAATCAGCTCTGAGACTGGGTTA	2640
QY	2641	GAATGTAAACAGCTTTAACTTGGGAAATTTAAGAGCTTTTAAAGGTAAATATCTCTGAA	2700
Db	2641	GAATGTAAACAGCTTTAACTTGGGAAATTTAAGAGCTTTTAAAGGTAAATATCTCTGAA	2700
QY	2701	GAAAAATGACGTAAACCAAGCGTGTACTATGAAAGCTTATTTTAAATAAAGACGTGG	2760
Db	2701	GAAAAATGACGTAAACCAAGCGTGTACTATGAAAGCTTATTTTAAATAAAGACGTGG	2760
QY	2761	GCCATGAACTCATACCTGCCAATGAGTCAACATAGTATCTTTATGTATGATTAAGTT	2820
Db	2761	GCCATGAACTCATACCTGCCAATGAGTCAACATAGTATCTTTATGTATGATTAAGTT	2820
QY	2821	ACTAAATATATATTTCACTTACTCTGAGGTGATAGTCTTCCCTCCCTCCCACTTTT	2880
Db	2821	ACTAAATATATATTTCACTTACTCTGAGGTGATAGTCTTCCCTCCCTCCCACTTTT	2880
QY	2881	TCTTTTGAAGGCAAGTGTGATCACTGAGGCGAGAGTGTGAGACCAAGCTGGGCAAT	2940
Db	2881	TCTTTTGAAGGCAAGTGTGATCACTGAGGCGAGAGTGTGAGACCAAGCTGGGCAAT	2940
QY	2941	AGCGAAACCCGATCTCTCTAATAAAAAATTTAAAAATTGGCCGAGCTGTGGCATGCT	3000
Db	2941	AGCGAAACCCGATCTCTCTAATAAAAAATTTAAAAATTGGCCGAGCATGTGGCATGCT	3000
QY	3001	GTGTCTCCAGCTACTCGGGAGGTTGAGGCAAGAGAGTGTGTAATGACGAGGTGAGG	3060
Db	3001	GTGTCTCCAGCTACTCGGGAGGTTGAGGCAAGAGAGTGTGTAATGACGAGGTGAGG	3060
QY	3061	TTGCAATGAGCAAGATTGTGCACCTGCACAGCTGGGGCAACAGACAAAGTCTGTC	3120
Db	3061	TTGCAATGAGCAAGATTGTGCACCTGCACAGCTGGGGCAACAGACAAAGTCTGTC	3120
QY	3121	TC 3122	
Db	3121	TC 3122	
RESULT 3			
AC008736/c			
LOCUS	AC008736	191925 bp	DNA linear PRI 27-SEP-2000
DEFINITION	Homo sapiens chromosome 19 clone CTD-2538C1, complete sequence.		
ACCESSION	AC008736		
VERSION	AC008736.6	GI:10312244	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Homnidae; Homo.		
REFERENCE	1 (bases 1 to 191925)		
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.		
TITLE	Direct Submission		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 191925)		
AUTHORS	DOE Joint Genome Institute.		
TITLE	Direct Submission		

JOURNAL	Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA				
REFERENCE	3 (bases 1 to 191925)				
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.				
TITLE	Direct Submission				
JOURNAL	Submitted (27-SEP-2000) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA				
COMMENT	On Sep 27, 2000 this sequence version replaced gi:8575905. Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov Finishing Completed at Stanford Human Genome Center www.sngc.stanford.edu Quality: Phrap Quality >=40 99.9% of Sequence; Estimated Total Number of Errors is 0.1. STS Content: SHGC-57769 G37408.				
FEATURES	Location/Qualifiers				
Source	1. 191925 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="19" /clone="CTD-2538C1"				
ORIGIN					
Query Match	91.0%	Score 2840;	DB 8;	Length 191925;	
Best Local Similarity	99.9%;	Pred. No. 0;			
Matches 2890;	Conservative	0;	Mismatches	1;	Indels 0; Gaps 0;
QY	1 ACTAGAGGTGGGGTTAGAGGCTTGGAAAGCACGCAACGATGAGCGCAACGCGGACAGGAGC 60				
Db	129460 ACTAGAGGTGGGGTTAGAGGCTTGGAAAGCACGCAACGATGAGCGCAACGCGGACAGGAGC 129401				
QY	61 AACTGACCCCGCGCGCGCCGACGCCCTCGATTTGCACTGCTCGCTTTGCGGACAG 120				
Db	129400 AACTGACCCCGCGCGCGCCGACGCCCTCGATTTGCACTGCTCGCTTTGCGGACAG 129341				
QY	121 GAGGTGCCCACTCTCGCGGGGACCCCGACGTCCTGTCGCGGACAGAGTCCGGAGTCACT 180				
Db	129340 GAGGTGCCCACTCTCGCGGGGACCCCGACGTCCTGTCGCGGACAGAGTCCGGAGTCACT 129281				
QY	181 ATAGCTGGGTTCTAGTCCATTCACAGGCAAAAACCTCCGCGGAGGCTGACCCTTTTAA 240				
Db	129280 ATAGCTGGGTTCTAGTCCATTCACAGGCAAAAACCTCCGCGGAGGCTGACCCTTTTAA 129221				
QY	241 CCTGGGCTCAAGTTCCCATCCGTAATATAGACGAGTTGATCTCCGAGCGCTAAC 300				
Db	129220 CCTGGGCTCAAGTTCCCATCCGTAATATAGACGAGTTGATCTCCGAGCGCTAAC 129161				
QY	301 TTCCGAACCTGGATGGGGCGAAGGGGAGGAGGGATGCGGCAACCCACAGTGAACCTCCC 360				
Db	129160 TTCCGAACCTGGATGGGGCGAAGGGGAGGAGGGATGCGGCAACCCACAGTGAACCTCCC 129101				
QY	361 CGCGTGAAGCCCGGCTCACTCACTGATTCAGAGGGGTGACAGCTCCGCGCGGAGCAGCGGG 420				
Db	129100 CGCGTGAAGCCCGGCTCACTCACTGATTCAGAGGGGTGACAGCTCCGCGCGGAGCAGCGGG 129041				
QY	421 GTGGCGGGGTCTTAGAAAACTTACCGGCGCGCTTGGCAACGCTTAAGGGGAGCGGG 480				
Db	129040 GTGGCGGGGTCTTAGAAAACTTACCGGCGCGCTTGGCAACGCTTAAGGGGAGCGGG 128981				
QY	481 CGGCTCTCAGCGCTCTTGCCCGGAGTTGGACCCACGAGAGATGGGGACGCACTTC 540				
Db	128980 CGGCTCTCAGCGCTCTTGCCCGGAGTTGGACCCACGAGAGATGGGGACGCACTTC 128921				
QY	541 AGCTTCGACGGAGCACCGTGGAGGCCAGGGCGGTGACAGACAGAGTGTGATCTCG 600				
Db	128920 AGCTTCGACGGAGCACCGTGGAGGCCAGGGCGGTGACAGACAGAGTGTGATCTCG 128861				
QY	601 AGTGGCGCTGGGGAGATGATGACAGAGGAGCGGGGAGCCGCTAACGGGGCTCCCTTGCGC 660				
Db	128860 AGTGGCGCTGGGGAGATGATGACAGAGGAGCGGGGAGCCGCTAACGGGGCTCCCTTGCGC 128801				



QY 661 GCCCGCTCCGACAGAGCGACGTCGAGGGTCCGCGGCGGGCTCCGTGACGTTGCGAGTA 720  
Db 128800 GCGCCGCTCCGACAGAGCGACGTCGAGGGTCCGCGGCGGGCTCCGTGACGTTGCGAGTA 128741  
QY 721 GCGCCGAGGAGTCAACGACCAATGAAAGACGTTCTGTCCGCGCGGCCAAGGCCGAGATG 780  
Db 128740 GCGCCGAGGAGTCAACGACCAATGAAAGACGTTCTGTCCGCGCGGCCAAGGCCGAGATG 128681  
QY 781 GGGGTTAAGCAATCCCTGCGCGCTGAGGGGAGGAGGCTTAACGCGCGCGCGCGCGCGCGCG 840  
Db 128680 GGGGTTAAGCAATCCCTGCGCGCTGAGGGGAGGAGGCTTAACGCGCGCGCGCGCGCGCGCG 128621  
QY 841 ACCCGAGGCCACCGCGATGCGAGGGGAGAGTCAAGGCGCTGTGACGCGGCTCAACA 900  
Db 128620 ACCCGAGGCCACCGCGAGTGGCGAGGGAGAGTCAAGGCGCTGTGACGCGGCTCAACA 128561  
QY 901 AGACGACTGCGCTGACCAACCTGTGTCTGACCGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960  
Db 128560 AGACGACTGCGCTGACCAACCTGTGTCTGACCGTGTGTGTGTGTGTGTGTGTGTGTGTGT 128501  
QY 961 ACCCTGCGGAGAGCTGCAAAAACGCGCGCAAGAGCGCGAGAGCTGTGTGTGTGTGTGTGTGT 1020  
Db 128500 ACCCTGCGGAGAGCTGCAAAAACGCGCGCAAGAGCGCGAGAGCTGTGTGTGTGTGTGTGTGT 128441  
QY 1021 GCGCCGCGCTGACTGT 1080  
Db 128440 GCGCCGCGCTGACTGT 128381  
QY 1081 TCGAGCGGCTGT 1140  
Db 128380 TCGAGCGGCTGT 128321  
QY 1141 GCTCGCTGAGAGCTGT 1200  
Db 128320 GCTCGCTGAGAGCTGT 128261  
QY 1201 CAGGT 1260  
Db 128260 CAGGT 128201  
QY 1261 GCGCTGAGGCGGAGGGCGACTTTCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1320  
Db 128200 GCGCTGAGGCGGAGGGCGACTTTCGAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 128141  
QY 1321 TTCAGGT 1380  
Db 128140 TTCAGGT 128081  
QY 1381 TCGAAGCCCGGAGGGCGGGGCGCGAGGCTCTGTTCACGAGTCAAGCGCGCGCGCTCTCTCT 1440  
Db 128080 TCGAAGCCCGGAGGGCGGGGCGCGAGGCTCTGTTCACGAGTCAAGCGCGCGCGCTCTCTCT 128021  
QY 1441 CCGTCTGTCTCTTCAGAGCGCGGGGGGGGTTTCAGACCCAGAGAGGCCCTGTGCGCGCA 1500  
Db 128020 CCGTCTGTCTCTTCAGAGCGCGGGGGGGGTTTCAGACCCAGAGAGGCCCTGTGCGCGCA 127961  
QY 1501 TCCCTTTTCCGCGCGCTGT 1560  
Db 127960 TCCCTTTTCCGCGCGCTGT 127901  
QY 1561 GCTGACAGACACCGGAGCGCGCTGT 1620  
Db 127900 GCTGACAGACACCGGAGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 127841  
QY 1621 GATGT 1680  
Db 127840 GATGT 127781  
QY 1681 GTCACAGCGCGCTTTCAGATGT 1740  
Db 127780 GTCACAGCGCGCTTTCAGATGT 127721  
QY 1741 CCGGAGAGAGTTAACTTTGCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1800

Db 127720 CCGGAGAGAGTTAACTTTGCGCGCGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 127661  
QY 1801 TTAATTCCTTAATTAATGAAAAACCTGCACATGTGATCCCTTCCGAGTTAATGAGTT 1860  
Db 127660 TTAATTCCTTAATTAATGAAAAACCTGCACATGTGATCCCTTCCGAGTTAATGAGTT 127601  
QY 1861 AACACATGT 1920  
Db 127600 AACACATGT 127541  
QY 1921 GCGTCCCCCTTTCTGCGTGGGACAGTTTGAAGGTGGGTGGGTGGAGTGAAGTTTG 1980  
Db 127540 GCGTCCCCCTTTCTGCGTGGGACAGTTTGAAGGTGGGTGGGTGGAGTGAAGTTTG 127481  
QY 1981 GAGAGGAGACGCTGT 2040  
Db 127480 GAGAGGAGACGCTGT 127421  
QY 2041 TCAATGTCAAGAGCTTTTAATTAATCTTAATCTTTAAGGCGCTAATTTTGAAGAGTGTCC 2100  
Db 127420 TCAATGTCAAGAGCTTTTAATTAATCTTTAAGGCGCTAATTTTGAAGAGTGTGTCC 127361  
QY 2101 TGAGAGAGTTCAATCAAAAGGCTTTCTTAAGAGCGCTACAGCCCTTCTACAGAGT 2160  
Db 127360 TGAGAGAGTTCAATCAAAAGGCTTTCTTAAGAGCGCTACAGCCCTTCTACAGAGT 127301  
QY 2161 TTAATCAATTCCTTCCCAAGAGAGCTAAGAGAGTTTGAAGTCAATGACTTCCCACTGCCG 2220  
Db 127300 TTAATCAATTCCTTCCCAAGAGAGCTAAGAGAGTTTGAAGTCAATGACTTCCCACTGCCG 127241  
QY 2221 CTCAGGGGCTGACCTTAATTAAGAAAAACAAAGGGGTGGGTGAACCTTACTCTACAGAG 2280  
Db 127240 CTCAGGGGCTGACCTTAATTAAGAAAAACAAAGGGGTGGGTGAACCTTACTCTACAGAG 127181  
QY 2281 TTGATTCACATGT 2340  
Db 127180 TTGATTCACATGT 127121  
QY 2341 GGGTAAGAGAGAGACAGAGGCTTGGGGTGAAGGCCCACTGTGTGTGTGTGTGTGTGTGTGT 2400  
Db 127120 GGGTAAGAGAGAGACAGAGGCTTGGGGTGAAGGCCCACTGTGTGTGTGTGTGTGTGTGTGT 127061  
QY 2401 TCCCTCTGCGGGCTTAATTTTGTTCAGAACTTACAGAGTGTGTGAACCTCTTGTGA 2460  
Db 127060 TCCCTCTGCGGGCTTAATTTTGTTCAGAACTTACAGAGTGTGTGAACCTCTTGTGA 127001  
QY 2461 GAGGGGCTGGGAATCCTCTTTAGACACTTAATCCTAATTTATCCCTGGAATGTGTGTG 2520  
Db 127000 GAGGGGCTGGGAATCCTCTTTAGACACTTAATCCTAATTTATCCCTGGAATGTGTGTG 126941  
QY 2521 TGGCCAGTAGAGGCGCTTGTGGAGTCCCTGACCCCGCGCGTCCCGCCCTCCGCG 2580  
Db 126940 TGGCCAGTAGAGGCGCTTGTGGAGTCCCTGACCCCGCGCGTCCCGCCCTCCGCG 126881  
QY 2581 GGTAAATGTGGCATTTCTGGCCCAAGAGGTTTGAAGCAATGAGTCTGTGAGCTGGGTTA 2640  
Db 126880 GGTAAATGTGGCATTTCTGGCCCAAGAGGTTTGAAGCAATGAGTCTGTGAGCTGGGTTA 126821  
QY 2641 GAATGTAAACGCTTTAATCTTGTGATTTAAGAGCTTTTAAAGTTAATATCTGTGAA 2700  
Db 126820 GAATGTAAACGCTTTAATCTTGTGATTTAAGAGCTTTTAAAGTTAATATCTGTGAA 126761  
QY 2701 GAAAAATGACGTAAACACAGCGTGTACTATGAAAGCTGTATTTTAAATGAAGCGCTGG 2760  
Db 126760 GAAAAATGACGTAAACACAGCGTGTACTATGAAAGCTGTATTTTAAATGAAGCGCTGG 126701  
QY 2761 GGCATGAATCTATCTGCAATGTGATCAATGATGATGATGATGATGATGATGATGATGATG 2820  
Db 126700 GGCATGAATCTATCTGCAATGTGATCAATGATGATGATGATGATGATGATGATGATGATG 126641  
QY 2821 ACTAAATATATATTTCAATCTTGTGAAGTGTGATGATGATGATGATGATGATGATGATGAT 2880



Db 126640 ACTAATATATATTCATCTTGAAGTATAGTCTCCGCCGCCCACTTTT 126581  
QY 2881 TCTTTTTCAG 2891  
Db 126580 TCTTTTTCAG 126570

RESULT 4  
AC027340/c 216441 bp DNA 1linear HTG 15-UTL-2000  
LOCUS Homo sapiens chromosome 19 clone CTD-2257C19, WORKING DRAFT  
DEFINITION SEQUENCE, 52 unordered pieces.  
AC027340  
AC027340.2 GI:9211228  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
1 (bases 1 to 216441)  
DOE Joint Genome Institute.  
Sequencing of Human Chromosome 19  
Unpublished  
2 (bases 1 to 216441)  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (30-MAR-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
On Jul 15, 2000 this sequence version replaced gi:7341654.  
-----Genome Center  
Center: Joint Genome Institute  
Web site: <http://www.jgi.doe.gov>  
-----  
Project Information  
Center Project Name: 726304, BC691328  
Center clone name: CITB-HI\_2257C19  
-----

Summary Statistics  
Consensus quality: 164680 bases at least Q40  
Consensus quality: 187929 bases at least Q30  
Consensus quality: 195340 bases at least Q20  
Estimated insert size: 159060; agarose-fp estimation  
Estimated insert size: 211341; sum-of-contigs estimation  
Quality coverage: 8.59 in Q20 bases; agarose-fp estimation  
Quality coverage: 6.46 in Q20 bases; sum-of-contigs estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 52 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1041: contig of 1041 bp in length  
\* 1042 1141: gap of unknown length  
\* 1142 2370: contig of 1229 bp in length  
\* 2371 2470: gap of unknown length  
\* 2471 3552: contig of 1082 bp in length  
\* 3553 3652: gap of unknown length  
\* 3653 4683: contig of 1031 bp in length  
\* 4684 4783: gap of unknown length  
\* 4784 6178: contig of 1395 bp in length  
\* 6179 6278: gap of unknown length  
\* 6279 7442: contig of 1164 bp in length  
\* 7443 7542: gap of unknown length  
\* 7543 8661: contig of 1119 bp in length  
\* 8662 8761: gap of unknown length  
\* 8762 10030: contig of 1269 bp in length  
\* 10031 10130: gap of unknown length  
\* 10131 11491: contig of 1361 bp in length  
\* 11492 11591: gap of unknown length

\* 11592 12967: contig of 1376 bp in length  
\* 12968 13067: gap of unknown length  
\* 13068 14173: contig of 1106 bp in length  
\* 14174 14273: gap of unknown length  
\* 14274 15840: contig of 1567 bp in length  
\* 15841 15940: gap of unknown length  
\* 15941 17242: contig of 1302 bp in length  
\* 17243 17342: gap of unknown length  
\* 17343 18449: contig of 1107 bp in length  
\* 18450 18549: gap of unknown length  
\* 18550 19683: contig of 1134 bp in length  
\* 19684 21247: gap of unknown length  
\* 19784 21347: contig of 1464 bp in length  
\* 21348 22952: gap of unknown length  
\* 21349 23052: contig of 1605 bp in length  
\* 22953 23052: gap of unknown length  
\* 23053 24452: contig of 1400 bp in length  
\* 24453 24552: gap of unknown length  
\* 24553 25825: contig of 1273 bp in length  
\* 25826 25925: gap of unknown length  
\* 25926 27109: contig of 1184 bp in length  
\* 27110 27209: gap of unknown length  
\* 27210 28292: contig of 1083 bp in length  
\* 28293 28392: gap of unknown length  
\* 28393 29744: contig of 1352 bp in length  
\* 29745 29844: gap of unknown length  
\* 29845 31042: contig of 1198 bp in length  
\* 31043 31142: gap of unknown length  
\* 31143 32629: contig of 1487 bp in length  
\* 31143 32629: gap of unknown length  
\* 32630 33279: gap of unknown length  
\* 33279 33898: contig of 1169 bp in length  
\* 33899 33998: gap of unknown length  
\* 33999 35704: contig of 1706 bp in length  
\* 35705 35804: gap of unknown length  
\* 35805 37471: contig of 1667 bp in length  
\* 37472 37571: gap of unknown length  
\* 37572 39285: contig of 1714 bp in length  
\* 39286 39385: gap of unknown length  
\* 39386 40577: contig of 1192 bp in length  
\* 40578 40677: gap of unknown length  
\* 40678 41930: contig of 1253 bp in length  
\* 41931 42030: gap of unknown length  
\* 42031 43743: contig of 1713 bp in length  
\* 43744 43843: gap of unknown length  
\* 43844 46054: contig of 2211 bp in length  
\* 46055 46154: gap of unknown length  
\* 46155 49379: contig of 3225 bp in length  
\* 49379 49479: gap of unknown length  
\* 49480 53371: contig of 3892 bp in length  
\* 53372 53471: gap of unknown length  
\* 53472 55756: contig of 2285 bp in length  
\* 55757 55856: gap of unknown length  
\* 55857 58019: contig of 2163 bp in length  
\* 58019 58119: gap of unknown length  
\* 58120 60327: contig of 2208 bp in length  
\* 60327 60427: gap of unknown length  
\* 60428 64323: contig of 3796 bp in length  
\* 64324 64323: gap of unknown length  
\* 64324 67218: contig of 2885 bp in length  
\* 67219 67318: gap of unknown length  
\* 67319 71797: contig of 4479 bp in length  
\* 71798 71897: gap of unknown length  
\* 71898 76856: contig of 4959 bp in length  
\* 76857 76956: gap of unknown length  
\* 76957 82060: contig of 5104 bp in length  
\* 82061 82160: gap of unknown length  
\* 82161 87557: contig of 5397 bp in length  
\* 87558 87657: gap of unknown length  
\* 87658 94326: contig of 6569 bp in length  
\* 94327 101878: gap of unknown length  
\* 101879 101978: contig of 7552 bp in length  
\* 101979 110629: contig of 8651 bp in length

\* 110630 110729: gap of unknown length  
\* 110730 119083: contig of 8354 bp in length  
\* 119084 119183: gap of unknown length  
\* 119184 128919: contig of 9736 bp in length  
\* 128920 129019: gap of unknown length  
\* 129020 142791: contig of 13772 bp in length  
\* 142792 142891: gap of unknown length  
\* 142892 156079: contig of 13188 bp in length  
\* 156080 156179: gap of unknown length  
\* 156180 183092: contig of 26913 bp in length  
\* 183093 183192: gap of unknown length  
\* 183193 216441: contig of 33249 bp in length.  
Location/Qualifiers  
1. 216441

## FEATURES

source

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clone="CTD-2257C19"  
/clone\_1lb="Caltech human BAC library D"  
1042..1141  
gap /estimated\_length=unknown  
2371..2470  
gap /estimated\_length=unknown  
3553..3652  
gap /estimated\_length=unknown  
4684..4783  
gap /estimated\_length=unknown  
6179..6278  
gap /estimated\_length=unknown  
7443..7542  
gap /estimated\_length=unknown  
8662..8761  
gap /estimated\_length=unknown  
10031..10130  
gap /estimated\_length=unknown  
11492..11591  
gap /estimated\_length=unknown  
12968..13067  
gap /estimated\_length=unknown  
14174..14273  
gap /estimated\_length=unknown  
15841..15940  
gap /estimated\_length=unknown  
17243..17342  
gap /estimated\_length=unknown  
18450..18549  
gap /estimated\_length=unknown  
19684..19783  
gap /estimated\_length=unknown  
21248..21347  
gap /estimated\_length=unknown  
22953..23052  
gap /estimated\_length=unknown  
24453..24552  
gap /estimated\_length=unknown  
25826..25925  
gap /estimated\_length=unknown  
27110..27209  
gap /estimated\_length=unknown  
28293..28392  
gap /estimated\_length=unknown  
29745..29844  
gap

Query Match 91.0%; Score 2840; DB 14; Length 216441;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 2890; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ACTAGAGGTGGGTTGCGCTTGAAGCACCAGCAGTGAAGCGGACGCGGAGGAC 60  
DB 117034 ACTAGAGGTGGGTTGCGCTTGAAGCACCAGCAGTGAAGCGGACGCGGAGGAC 116975  
QY 61 ACCTGACCCGCGCGCCAGCCCTCGGATTGCCAGTCACTGCTGCTTGGGCGACG 120

DB 116974 ACCTGACCCGCGCGCCAGCCCTCGGATTGCCAGTCACTGCTGCTTGGGCGACG 116915  
QY 121 GAGGTGCCCAAGTCTCGCGGGGACCCGACGTCCTTCCGCCACAGGCTCCGGAGTCACT 180  
DB 116914 GAGGTGCCCAAGTCTCGCGGGGACCCGACGTCCTTCCGCCACAGGCTCCGGAGTCACT 116855  
QY 181 ATAGCTGGGTTTGTAGTCCATCAAGGCAAAAATCCGCGGGAGGCTGGCCCGCTTTTGA 240  
DB 116854 ATAGCTGGGTTTGTAGTCCATCAAGGCAAAAATCCGCGGGAGGCTGGCCCGCTTTTGA 116795  
QY 241 CCTGGGCTCAAGTTTCCCATTCGTAATAATAGAACGCGTTGATCTCCGAGCGCTTAA 300  
DB 116794 CCGGAGGCTCAAGTTTCCCATTCGTAATAATAGAACGCGTTGATCTCCGAGCGCTTAA 116735  
QY 301 TTCCGAACCTCGGATGGGGGCGAAGGGAGGGAGTGGGCCACCAACAGTCACTCC 360  
DB 116734 TTCCGAACCTCGGATGGGGGCGAAGGGAGGGAGTGGGCCACCAACAGTCACTCC 116675  
QY 361 CGCGTGAAGCCCGGCTTCACTGATCCAGGGGGTGGAGCTCCGCGCGGAGCGAGCGG 420  
DB 116674 CGCGTGAAGCCCGGCTTCACTGATCCAGGGGGTGGAGCTCCGCGCGGAGCGAGCGG 116615  
QY 421 GTGGCGGGTCTCTAGAAACCTTACC CGGCGCTTGGCAAGCGCTTAAAGCGAGCGCG 480  
DB 116614 GTGGCGGGTCTCTAGAAACCTTACC CGGCGCTTGGCAAGCGCTTAAAGCGAGCGCG 116555  
QY 481 CGGCTCTGAGCTCTGCTTCCCGGAGTTTGGACCAACGAGAGATGGGGACCGCACCTTC 540  
DB 116554 CGGCTCTGAGCTCTGCTTCCCGGAGTTTGGACCAACGAGAGATGGGGACCGCACCTTC 116495  
QY 541 AGCTTGAGAGGAGCACCCTGAGAGGCGAGGCGGTGAGAGACAGACGTTGACTCGG 600  
DB 116494 AGCTTGAGAGGAGCACCCTGAGAGGCGAGGCGGTGAGAGACAGACGTTGACTCGG 116435  
QY 601 AGTGGCGCTTGGAGAGATGAGACGAGGAGCGGGAGACCGTTAACGGGCTCCCTTGGCG 660  
DB 116434 AGTGGCGCTTGGAGAGATGAGACGAGGAGCGGGAGACCGTTAACGGGCTCCCTTGGCG 116375  
QY 661 GCCCGCTCGCAGAGAGCGCAGTCCAGAGGCTCCCGGGCGGGCTCCGTGACAGTTGGCGGTA 720  
DB 116374 GCCCGCTCGCAGAGAGCGCAGTCCAGAGGCTCCCGGGCGGGCTCCGTGACAGTTGGCGGTA 116315  
QY 721 GCGCGAGAGGTCAACGAGCAATGAAGCGTTGTCGTCGCGCGGCCACAGGCGGGAGTG 780  
DB 116314 GCGCGAGAGGTCAACGAGCAATGAAGCGTTGTCGTCGCGCGGCCACAGGCGGGAGTG 116255  
QY 781 GGGGTTTAACTATCTGCGCGCTGAGAGGAGGCTTAAACGCGCGCGGCGCGGCGCC 840  
DB 116254 GGGGTTTAACTATCTGCGCGCTGAGAGGAGGCTTAAACGCGCGCGGCGCGGCGCC 116195  
QY 841 AGCCGAGGCCACCGCGATGCGAGAGAGAGAGTGAAGGCGCTGCTGACAGGGCTTCAACA 900  
DB 116194 AGCCGAGGCCACCGCGATGCGAGAGAGAGAGTGAAGGCGCTGCTGACAGGGCTTCAACA 116135  
QY 901 AGACGACTGCGTGTACCAACCACTGATGAGACCGTGGTGGTCTCGGCGACTCGCAGA 960  
DB 116134 AGACGACTGCGTGTACCAACCACTGATGAGACCGTGGTGGTCTCGGCGACTCGCAGA 116075  
QY 961 ACTTCGCGCAGAGCTGCAAAAAGACGCGCAGAAAGCGCAGAGAGCTGCGGTGCCACT 1020  
DB 116074 ACTTCGCGCAGAGCTGCAAAAAGACGCGCAGAAAGCGCAGAGAGCTGCGGTGCCACT 116015  
QY 1021 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCTTGGCGCGACAGAGCGCGCGAGT 1080  
DB 116014 GCGCCCGGCTGACTGCTGTGCTGCGCGACCGGGGCTTGGCGCGACAGAGCGCGCGAGT 115955  
QY 1081 TCGAGCGGCTCTGGGTGGGCTTCTCGGGCTGCTGACCTGCTGGAAGCGGACATGCGAC 1140  
DB 115954 TCGAGCGGCTCTGGGTGGGCTTCTCGGGCTGCTGACCTGCTGGAAGCGGACATGCGAC 115895  
QY 1141 GCTGCTGAGAGCTGGGGCGCGGCTTCCGCTGCAAGCGCGCGGAGACGCTGTGTCGCA 1200

Db 115894 GCTCGGTGAGAGTGGGCGCGCGCTTCCGCTGACGCGCGCGCGCGCTGATGCGCA 115835  
Qy 1201 CAGGTGTGCTGCGCGCTCTCTCCGCGTGGCGCGCGCGCTGAGCACCCGCAAGCTGC 1260  
Db 115834 CAGGTGTGCTGCGCGCTCTCTCCGCGTGGCGCGCGCGCTGAGCACCCGCAAGCTGC 115775  
Qy 1261 GGTCTGAGGCGGAGGCGGCGCTTCCGAGTGGCGGACCTTGGGAGAGCTGAGCGCGAGTGC 1320  
Db 115774 GGTCTGAGGCGGAGGCGGAGCTTCCGAGTGGCGGACCTTGGGAGAGCTGAGCGCGAGTGC 115715  
Qy 1321 TTCAAGTGGGCGAGATGATCGACACATGAGATGAGAGTCAACGTGCCCGCTGAGACCG 1380  
Db 115714 TTCAAGTGGGCGAGATGATCGACACATGAGATGAGAGTCAACGTGCCCGCTGAGACCG 115655  
Qy 1381 TGGCAAGCCGCGGAGGCGGCGCGCGAGCTCTGTGCAACGCTCAGCGCGCGCGCTCTCT 1440  
Db 115654 TGGCAAGCCGCGGAGGCGGCGCGCGAGCTCTGTGCAACGCTCAGCGCGCGCGCTCTCT 115595  
Qy 1441 CGGTGCTGCTTGGAGAGCGCGGCGGAGTGGGAGACCCGAGAGAGCGCTGGCGCGCA 1500  
Db 115594 CGGTGCTGCTTGGAGAGCGCGGCGGAGTGGGAGACCCGAGAGAGCGCTGGCGCGCA 115535  
Qy 1501 TCTTTTCGCGCGCGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1560  
Db 115534 TCTTTTCGCGCGCGCTGCTGCTGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 115475  
Qy 1561 GCTGACAGACACCCGACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1620  
Db 115474 GCTGACAGACACCCGACGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 115415  
Qy 1621 GATGGGTGGGGGTCTGCGCTGCGCAAGGGGAGTGGTCTTAAACCCGCTGCTGCTGCTG 1680  
Db 115414 GATGGGTGGGGGTCTGCGCTGCGCAAGGGGAGTGGTCTTAAACCCGCTGCTGCTGCTG 115355  
Qy 1681 GTACACGCGCGCTTCCAGTGCATCTGCTGCGGAGAGCAACGCTTTTCTCTTGTGCTGC 1740  
Db 115354 GTACACGCGCGCTTCCAGTGCATCTGCTGCGGAGAGCAACGCTTTTCTCTTGTGCTGC 115295  
Qy 1741 CCGGGAGAGTTACCTTTGCGCGCGCGCTGACGGGCACTTACCGCTTACGCTGACAGAGCT 1800  
Db 115294 CCGGGAGAGTTACCTTTGCGCGCGCGCTGACGGGCACTTACCGCTTACGCTGACAGAGCT 115235  
Qy 1801 TTATTCCTATTAAATAGAAAACCGTCAAGTGAACCTAGATCCCTGCGAGTTAATAGTT 1860  
Db 115234 TTATTCCTATTAAATAGAAAACCGTCAAGTGAACCTAGATCCCTGCGAGTTAATAGTT 115175  
Qy 1861 AACACATGTGCTGTGGGCGCTTTCACAGGAGTCCGAGTTGCGTCCACCCCTGCCA 1920  
Db 115174 AACACATGTGCTGTGGGCGCTTTCACAGGAGTCCGAGTTGCGTCCACCCCTGCCA 115115  
Qy 1921 GCGTCCCGCTTTCGCGTGGGAGAGTTTGAAGAGTGGGTGGAGTGAAGTTTG 1980  
Db 115114 GCGTCCCGCTTTCGCGTGGGAGAGTTTGAAGAGTGGGTGGAGTGAAGTTTG 115055  
Qy 1981 GAGAGGAGCGCTGTTGTTCTATGAGTGTGCTGTTTCCGCGACAGAAAAATTGCA 2040  
Db 115054 GAGAGGAGCGCTGTTGTTCTATGAGTGTGCTGTTTCCGCGACAGAAAAATTGCA 114995  
Qy 2041 TCAAAATGTGAGAGCTTTTATTAACCTTAACTTTTCAAGGCGCTTAAATTTAGAGAGTGC 2100  
Db 114994 TCAAAATGTGAGAGCTTTTATTAACCTTAACTTTTCAAGGCGCTTAAATTTAGAGAGTGC 114935  
Qy 2101 TGAAGACATTCATTAAGAGGCTTCTCTAAGACCGCTTACAGCCCTTCTAGCAGAT 2160  
Db 114934 TGAAGACATTCATTAAGAGGCTTCTCTAAGACCGCTTACAGCCCTTCTAGCAGAT 114875  
Qy 2161 TTATTCATTCGTCGCCAAGAGAGCTAGAAAGATTTGAGGTGATGACCTCCCACTGCGG 2220  
Db 114874 TTATTCATTCGTCGCCAAGAGAGCTAGAAAGATTTGAGGTGATGACCTCCCACTGCGG 114815  
Qy 2221 CTCAGGGGCTGACCTCTATTAGAGAAACAAAGAGGCTGGGTGGAACCTACTCAGAGAC 2280  
Db 114814 CTCAGGGGCTGACCTCTATTAGAGAAACAAAGAGGCTGGGTGGAACCTACTCAGAGAC 114755

Qy 2281 TTGATTCAGTCCGACACTTGCCTGCGGAAAAAGGCTCTCCCGACGCCACCGGAGATGG 2340  
Db 114754 TTGATTCAGTCCGACACTTGCCTGCGGAAAAAGGCTCTCCCGACGCCACCGGAGATGG 114695  
Qy 2341 GGGTAAGAGAGACAGAGGCTTGGGATGAGGCGCACCTGCTGTTTAAACAGGACCTTTC 2400  
Db 114694 GGGTAAGAGAGACAGAGGCTTGGGATGAGGCGCACCTGCTGTTTAAACAGGACCTTTC 114635  
Qy 2401 TCTTCTCTGCGGCTTATTTTGTTCAGAACTAGACAGAGTGTGAACTCTCTTTGCA 2460  
Db 114634 TCTTCTCTGCGGCTTATTTTGTTCAGAACTAGACAGAGTGTGAACTCTCTTTGCA 114575  
Qy 2461 GGAAGGCTGGAAATCTCTTTAGAGACCTTAACTCTATTTATCCCTGGAAATGTCGTGC 2520  
Db 114574 GGAAGGCTGGGAATCTCTTTAGAGACCTTAACTCTATTTATCCCTGGAAATGTCGTGC 114515  
Qy 2521 TGGCAGTAGAGGCTGGCTTTGGCAGCTCCCTGACCCCGCGCTGCGCGCGCTCTCGG 2580  
Db 114514 TGGCAGTAGAGGCTGGCTTTGGCAGCTCCCTGACCCCGCGCTGCGCGCGCTCTCGG 114455  
Qy 2581 GGTAAATGTGGCATTACTGCGCCACAGAGCTTTTGAAGCAATCAGCTTGAAGCTGGTTA 2640  
Db 114454 GGTAAATGTGGCATTACTGCGCCACAGAGCTTTTGAAGCAATCAGCTTGAAGCTGGTTA 114395  
Qy 2641 GAATGTAAACGCTTAACTTGGGATTTAAGAGCTTTTAAAGATTAATCTCTGAAA 2700  
Db 114394 GAATGTAAACGCTTAACTTGGGATTTAAGAGCTTTTAAAGATTAATCTCTGAAA 114335  
Qy 2701 GAAAAATGACGTAAACAGAGCGTGTACTATGAAAGCTGTATTTTAAATGAAGACGTGG 2760  
Db 114334 GAAAAATGACGTAAACAGAGCGTGTACTATGAAAGCTGTATTTTAAATGAAGACGTGG 114275  
Qy 2761 GCCATGAATCTATACCTGCGCAATGAGTCAACATATGATCTTATATGATGATCTTAAATT 2820  
Db 114274 GCCATGAATCTATACCTGCGCAATGAGTCAACATATGATCTTATATGATGATCTTAAATT 114215  
Qy 2821 ACTAAATATATTTTCACTCTCTGAAAGTTGATGATCTTCCCGCGCGCGCGCACTTTT 2880  
Db 114214 ACTAAATATATTTTCACTCTCTGAAAGTTGATGATCTTCCCGCGCGCGCGCACTTTT 114155  
Qy 2881 TCTTTTGTGAG 2891  
Db 114154 TCTTTTGTGAG 114144  
  
RESULT 5  
AC008474 186115 bp DNA linear PRI 13-JUL-2002  
LOCUS AC008474  
DEFINITION Homo sapiens chromosome 19 clone CTC-379B2, complete sequence.  
ACCESSION AC008474  
VERSION AC008474.9 GI:21743729  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 186115)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
AUTHORS  
TITLE Direct Submision  
JOURNAL  
2 (bases 1 to 186115)  
DOE Joint Genome Institute.  
AUTHORS  
TITLE Direct Submision  
JOURNAL  
3 (bases 1 to 186115)  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE  
DOE Joint Genome Institute and Stanford Human Genome Center.  
AUTHORS  
TITLE Direct Submision  
JOURNAL  
Submitted (01-JUL-2000) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
REFERENCE  
4 (bases 1 to 186115)

AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (17-OCT-2001) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
REFERENCE 5 (bases 1 to 186115)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (13-JUL-2002) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
COMMENT On Jul 13, 2002 this sequence version replaced gi:16197759.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www-shgc.stanford.edu  
Quality: Phrap Quality >=40 99.8% of Sequence;  
Estimated Total Number of Errors is 0.4.

FEATURES  
source  
1..186115  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clone="CTC-379B2"

ORIGIN

Query Match 78.3%; Score 2445; DB 8; Length 186115;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 2865; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

QY 1 ACTAGAGTGGGTTAGCGCTTGAGAGCACCGACCACTGAGCGGACGCGGAGGAC 60  
Db 180657 ACTAGAGTGGGTTAGCGCTTGAGAGCACCGACCACTGAGCGGACGCGGAGGAC 180716  
QY 61 ACCTGACCCCGCGCGCGCCAGCGCCCTCGAATTGCCAGTCACTGCTCGCTTTGGGACAG 120  
Db 180717 ACCTGACCCCGCGCGCGCCAGCGCCCTCGAATTGCCAGTCACTGCTCGCTTTGGGACAG 180776  
QY 121 GAGGAGCCAGTCTCGCGGGGAGACCGGAGTCTGTCGCGGACAGGAGTCCGGGAGTCACT 180  
Db 180777 GAGGAGCCAGTCTCGCGGGGAGACCGGAGTCTGTCGCGGACAGGAGTCCGGGAGTCACT 180836  
QY 181 ATAGCTGAGTCTAGTCCCATCAAGAGCAAAACTCCGCGGAGGCTGCGCCGCTTTTAA 240  
Db 180837 ATAGCTGAGTCTAGTCCCATCAAGAGCAAAACTCCGCGGAGGCTGCGCCGCTTTTAA 180896  
QY 241 CCTGGGCTCAAGTTTCCCATCCGTAAATAGACGGGTTGATCTCCGAGCGCTAAACA 300  
Db 180897 CCTGGGCTCAAGTTTCCCATCCGTAAATAGACGGGTTGATCTCCGAGCGCTAAACA 180956  
QY 301 TTCCGAAGTCCGAGTGGGCGGAGGAGGAGGAGTGGCCACCCACAGTGAACCTCCG 360  
Db 180957 TTCCGAAGTCCGAGTGGGCGGAGGAGGAGGAGTGGCCACCCACAGTGAACCTCCG 181016  
QY 361 CGCGTGAAGCCCGGCTTACCACTGATCAGAGGGGTGGCAGCTCCGCGCGGAGAGCGGG 420  
Db 181017 CGCGTGAAGCCCGGCTTACCACTGATCAGAGGGGTGGCAGCTCCGCGCGGAGAGCGGG 181076  
QY 421 GTGGGCGGGTCTTAGGAAACCTTACCAGGCGCGCTTGGCAGCGCTTAAAGCGGAGCGG 480  
Db 181077 GTGGGCGGGTCTTAGGAAACCTTACCAGGCGCGCTTGGCAGCGCTTAAAGCGGAGCGG 181136  
QY 481 CGGCTCTGAGCGCTGTTGCCCCCGAGTTGGACCCACGAGAGATGGGAGCCGCAACCTTC 540  
Db 181137 CGGCTCTGAGCGCTGTTGCCCCCGAGTTGGACCCACGAGAGATGGGAGCCGCAACCTTC 181196  
QY 541 AGCTTGCAGAGGACCAAGTGAAGGCGAGGGCGGTGCAGAGACAGAGTGTGACTCGG 600  
Db 181197 AGCTTGCAGAGGACCAAGTGAAGGCGAGGGCGGTGCAGAGACAGAGTGTGACTCGG 181256  
QY 601 AGTGCCTCTGGGAGATGAGCAGAGGAGCGGGGAGCCGCTTAAAGGGGCTCCCTCTGCGC 660  
Db 181257 AGTGCCTCTGGGAGATGAGCAGAGGAGCGGGGAGCCGCTTAAAGGGGCTCCCTCTGCGC 181316

QY 661 GCCCGCTCCAGAGAGCCGACGTCGAGGGTCCCGGCGGAGCTCCGTGACGTTGCGGTA 720  
Db 181317 GCCCGCTCCAGAGAGCCGACGTCGAGGGTCCCGGCGGAGCTCCGTGACGTTGCGGTA 181376  
QY 721 GCGCCGACGAGTCAAGGACATGAAGAGCTTGTGCTCGCGCGGCTCCAGAGCCGGAGT 780  
Db 181377 GCGCCGACGAGTCAAGGACATGAAGAGCTTGTGCTCGCGCGGCTCCAGAGCCGGAGT 181436  
QY 781 GGGGTTAACCATCTCTGCCCCGCTGAGGGGAGGCTTAAAGGCGCGGCGCGGCGGCG 840  
Db 181437 GGGGTTAACCATCTCTGCCCCGCTGAGGGGAGGCTTAAAGGCGCGGCGCGGCGGCG 181496  
QY 841 AGCCGAGACCCACCGGATGAGGAGGAGAGAGTGAAGGCGCTGTGTCAGAGCGGCTCAACA 900  
Db 181497 AGCCGAGACCCACCGGATGAGGAGGAGAGAGTGAAGGCGCTGTGTCAGAGCGGCTCAACA 181556  
QY 901 AGACGACTGCGTCAACCAACGCTGTGCTGACCGTGTGCTGTGCGGAGCTGCGCAGA 960  
Db 181557 AGACGACTGCGTCAACCAACGCTGTGCTGACCGTGTGCTGTGCGGAGCTGCGCAGA 181616  
QY 961 ACCTGCGGACGAGAGCTGCAAAAGAGCGCCGACAGAGAGGCGAGGCGGTGCACCT 1020  
Db 181617 ACCTGCGGACGAGAGCTGCAAAAGAGCGCCGACAGAGAGGCGAGGCGGTGCACCT 181676  
QY 1021 GCGCCGAGCTGACTGTGCTGCGGACCGGAGGCTGTGCGCGACGAGCGCGGAGT 1080  
Db 181677 GCGCCGAGCTGACTGTGCTGCGGACCGGAGGCTGTGCGCGACGAGCGCGGAGT 181736  
QY 1081 TCGAGCGGCTGTGGGTGAGCTTTCTGCGGCTGCTGAGACTGTGTAAGCGGACATGCGAC 1140  
Db 181737 TCGAGCGGCTGTGGGTGAGCTTTCTGCGGCTGCTGAGACTGTGTAAGCGGACATGCGAC 181796  
QY 1141 GCTCGCTGAGTGTGGGCGCGCGGCTTCCGCTGCAAGCGCCGCGGAGACCGCTGTGCGGA 1200  
Db 181797 GCTCGCTGAGTGTGGGCGCGCGGCTTCCGCTGCAAGCGCGCGGAGACCGCTGTGCGGA 181856  
QY 1201 CAGGTGTGAGCTGAGCGCTCTTCCGCGGTGAGCGCGCGCTGAGACCCGCAAGCTTC 1260  
Db 181857 CAGGTGTGAGCTGAGCGCTCTTCCGCGGTGAGCGCGCGCTGAGACCCGCAAGCTTC 181916  
QY 1261 GCGCTGAGCGGAGGAGGAGCTTTCAGAGTTCGAGGACCTGCGGAGGCTGAGAGCGGAGTTC 1320  
Db 181917 GCGCTGAGCGGAGGAGGAGCTTTCAGAGTTCGAGGACCTGCGGAGGCTGAGAGCGGAGTTC 181976  
QY 1321 TTCAAGTGGGAGATGATTCACCAATGAGAGATGAAGTCAACGTGCCCGCGTGAACCG 1380  
Db 181977 TTCAAGTGGGAGATGATTCACCAATGAGAGATGAAGTCAACGTGCCCGCGTGAACCG 182036  
QY 1381 TGCAGAGCCGAGGAGGCGGCGGCGGAGCTCTGTCAAGGTCAGCGCGGCGCCCTCTCT 1440  
Db 182037 TGCAGAGCCGAGGAGGCGGCGGCGGAGCTCTGTCAAGGTCAGCGCGGCGCCCTCTCT 182096  
QY 1441 CGGTGTGTCTTTCAGAGAGCGCGGAGGAGTTCGACCCCAAGAGGCTTGCCTGCCA 1500  
Db 182097 CGGTGTGTCTTTCAGAGAGCGCGGAGGAGTTCGACCCCAAGAGGCTTGCCTGCCA 182156  
QY 1501 TCTTTTGGGCGCGCTGTGCTGAGCGCTGTGAGCCCTTGAAGCGTGTGCGGAGAGCTGA 1560  
Db 182157 TCTTTTGGGCGCGCTGTGCTGAGCGCTGTGAGCCCTTGAAGCGTGTGCGGAGAGCTGA 182216  
QY 1561 GCTGACAGACCCGAGCGGCGCTGTGCTGCGCTCCCTCCCTGAGAAAGACTCGG 1620  
Db 182217 GCTGACAGACCCGAGCGGCGGCGCTGTGCTGCGCTCCCTCCCTGAGAAAGACTCGG 182276  
QY 1621 GATGGGTGTGGGTGTGAGCTGTGCAAGGAGATGTCTTAAACCCCGTGTGCATAG 1680  
Db 182277 GATGGGTGTGGGTGTGAGCTGTGCAAGGAGATGTCTTAAACCCCGTGTGCATAG 182336  
QY 1681 GTACACGGCGCTTTCAGATGACATCTGCTGTGGAGAGACAGGTTTCTCTTCTGCTGGC 1740  
Db 182337 GTACACGGCGCTTTCAGATGACATCTGCTGTGGAGAGACAGG-TTCTCTCTTCTGCTGGC 182395  
QY 1741 CGGAGGAAGTTTAACTTTGCGCGCGGCGGTCAAGGCAATTACCGCTTACGTCAGAGAGCT 1800

Db	182396	CGGAGGAGGATTACCTTGGCGCGCGCGCTCAGGGGATTACCCCTTAACGCTGACAGAGCT	182455
QY	1801	TTATTCCTTATTATAGAAAACCGTCACAGTACCCCTAGATCCCTCCGAGTTATGAGTT	1860
Db	182456	TTATTCCTTATTATAGAAAACCGTCACAGTACCCCTAGATCCCTCCGAGTTATGAGTT	182515
QY	1861	AACACATGCTGCTTGGGGCGCTTTTACAGGAGTCCGAGTTCCGATGCCACCCCTGCCA	1920
Db	182516	AACACATGCTGCTTGGGGCGCTTTTACAGGAGTCCGAGTTCCGATGCCACCCCTGCCA	182575
QY	1921	GGTCGCCCCCTTTCTGCGTGGGACAGTTGAAAAGTGGGTGGGGTGGAGTGAAGTTTG	1980
Db	182576	GGTCGCCCCCTTTCTGCGTGGGACAGTTGAAAAGTGGGTGGGGTGGAGTGAAGTTTG	182635
QY	1981	GAGAGGAGCGCTGTTGGTGTCTATGCTGTGTGTTCCCGACCAAGAAAATTGCA	2040
Db	182636	GAGAGGAGCGCTGTTGGTGTCTATGCTGTGTGTTCCCGACCAAGAAAATTGCA	182695
QY	2041	TCAATGTCAGCAGCTTTTATTAACCTTAATCTTTACAGGCGCTTAAATTTAGAGAGTGTCC	2100
Db	182696	TCAATGTCAGCAGCTTTTATTAACCTTAATCTTTACAGGCGCTTAAATTTAGAGAGTGTCC	182755
QY	2101	TGAGACAGCTTCAATACAAAGGCTTCTCTAAGAAGCGCTAGAGCCCTCTAGCAGAGT	2160
Db	182756	CGAGAGCATTCATACAAAGGCTTCTCTAAGAAGCGCTAGAGCCCTCTAGCAGAGT	182815
QY	2161	TTATTCATTGCTGCCCAAGAGAGCTAGAAAGATTGAGTCAATGCTCCCATCTGCCG	2220
Db	182816	TTATTCATTGCTGCCCAAGAGAGCTAGAAAGATTGAGTCAATGCTCCCATCTGCCG	182875
QY	2221	CTCAGGGGCTGACCCCTATTAGAAAACCAAGAGGTGGTGAACCTTCTCAGCAGAC	2280
Db	182876	CTCAGGGGCTGACCCCTATTAGAAAACCAAGAGGTGGTGAACCTTCTCAGCAGAC	182935
QY	2281	TTGATTCAGTGGGACACCTTGGCTGGGAAAAGGGCTCTCCCGACCCCGAGATGG	2340
Db	182936	TTGATTCAGTGGGACACCTTGGCTGGGAAAAGGGCTCTCCCGACCCCGAGATGG	182995
QY	2341	GGGTAAAGAGAAAGAGAGGCTGGGGGTAGGGCCACTGGGTGTTTAAACAGGCACTTTC	2400
Db	182996	GGGTAAAGAGAAAGAGAGGCTGGGGGTAGGGCCACTGGGTGTTTAAACAGGCACTTTC	183055
QY	2401	TGCTTCTCTGGGGCTTATTTTGTTCAGAACTAGACAGAGTGTGAACCTCTTGTGA	2460
Db	183056	TGCTTCTCTGGGGCTTATTTTGTTCAGAACTAGACAGAGTGTGAACCTCTTGTGA	183115
QY	2461	GGAGGCTGGGAATCCTCTTGTAGAGCACTTAATCTTAATTCCTCTGGAATGTGCTGC	2520
Db	183116	GGAGGCTGGGAATCCTCTTGTAGAGCACTTAATCTTAATTCCTCTGGAATGTGCTGC	183175
QY	2521	TGGCCAGTAGAGAGGCTGGCTTGGCACTCCCTGAACCCCGCGCTGCCCGCCCTCCG	2580
Db	183176	TGGCCAGTAGAGAGGCTGGCTTGGCACTCCCTGAACCCCGCGCTGCCCGCCCTCCG	183235
QY	2581	GGTAATGTGGCACTTACCTGGCCCAAGAGTTTGAACCACTAGAGCTGTGAGCTGGGTTA	2640
Db	183236	GGTAATGTGGCACTTACCTGGCCCAAGAGTTTGAACCACTAGAGCTGTGAGCTGGGTTA	183295
QY	2641	GAATGTACAGCTTTAACTTGGGATTTAAGAGCTTTTAAAGTAAATATCTCTGAAA	2700
Db	183296	GAATGTACAGCTTTAACTTGGGATTTAAGAGCTTTTAAAGTAAATATCTCTGAAA	183355
QY	2701	GAATAATGACGTACCAAGCGGTGACTATGAAAGCTGTTATTTTAAATAAGACGCTGG	2760
Db	183356	GAATAATGACGTACCAAGCGGTGACTATGAAAGCTGTTATTTTAAATAAGACGCTGG	183415
QY	2761	GCCATGAACTCAACCTGCGCAAGAGTCAAACTATATCTTATATGATCTTATGATTT	2820
Db	183416	GCCATGAACTCAACCTGCGCAAGAGTCAAACTATATCTTATATGATCTTATGATTT	183475
QY	2821	ACTAAATATATATTTCTACTTCTGAAGTTGATGCTTCCCCCCCCCCTTC	2872

Db	183476	ACTAAATATATTTCAATCTACTTCTGAAAGTTGATGATGCTTCCCCCCCCCTTC	183527		
RESULT 6	CO850499	2327 bp	DNA	linear	PAT 23-AUG-2004
LOCUS	Sequence 968 from Patent EP1447413.				
DEFINITION	CO850499				
ACCESSION	CO850499				
VERSION	CO850499.1	GI:51508711			
KEYWORDS					
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;				
	Homnidae; Homo.				
REFERENCE	1				
AUTHORS	Isegai,T., Yamamoto,J., Nishikawa,T., Isono,Y., Sugiyama,T.,				
	Otsuki,T., Makamatsu,A., Ishii,S., Nagai,K. and Irie,R.				
TITLE	Full-length human CDNA				
JOURNAL	Patent: EP 1447413-A 968 18-AUG-2004;				
FEATURES	Research Association for Biotechnology (JF)				
source	1..2327				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
ORIGIN					
Query Match	62.5%; Score 1952;	DB 6;	Length 2327;		
Best Local Similarity	99.7%; Pred. No. 0;				
Matches 2322;	Conservative 0;	Mismatches 5;	Indels 1;	Gaps 1;	
QY	444	ACCGGCGCGCCCTTGGGACGCGCTTAAGCGGAGCGCGGCTGACGCTTGGCCCC	503		
Db	1	ACCGGCGCGCCCTTGGGACGCGCTTAAGCGGAGCGCGGCTGACGCTTGGCCCC	60		
QY	504	GGAAGTTGGCAACCAACGAGATGGGAGACCGACCTCAGCTTGCAGAGGACCAACCGTGG	563		
Db	61	GGAAGTTGGCAACCAACGAGATGGGAGACCGACCTCAGCTTGCAGAGGACCAACCGTGG	120		
QY	564	AGGCCAGGGCGGTGCAGAGACAGACGCTGTGACTTGGAGTCCGCTGGGAGAGATGACG	623		
Db	121	AGGCCAGGGCGGTGCAGAGACAGACGCTGTGACTTGGAGTCCGCTGGGAGAGATGACG	180		
QY	624	AGGAGCGGGAGACCGCTTAACGGGCTCCCTGCGCGCCCGCTCCGAGAGGAGACGT	683		
Db	181	AGGAGCGGGAGACCGCTTAACGGGCTCCCTGCGCGCCCGCTCCGAGAGGAGACGT	240		
QY	684	CGAGGAGTCCCGGCGGCTCCGCTGAGAGCTTGGAGGTAGCGCCGAGAGTACAGACCAT	743		
Db	241	CGAGGAGTCCCGGCGGCTCCGCTGAGAGCTTGGAGGTAGCGCCGAGAGTACAGACCAT	300		
QY	744	GAAGAGCGTTGCTGTCGCGCGCGCCCAAGGCTGGAGTGGGGTTAGCCATCTGCGCG	803		
Db	301	GAAGAGCGTTGCTGTCGCGCGCGCCCAAGGCTGGAGTGGGGTTAGCCATCTGCGCG	360		
QY	804	CTGAGGGGAGGCTTAACGGGCGCGGCGCGCGCGCCAGCGGAGCCACCGGATGGCG	863		
Db	361	CTGAGGGGAGGCTTAACGGGCGCGGCGCGCGCGCCAGCGGAGCCACCGGATGGCG	420		
QY	864	AGGAGAGATGCAAGGCGCTGCTGAGACGGGCTCAACAGACGACTGTGTACACAC	923		
Db	421	AGGAGAGATGCAAGGCGCTGCTGAGACGGGCTCAACAGACGACTGTGTACACAC	480		
QY	924	CTGTGCTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	983		
Db	481	CTGTGCTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	540		
QY	984	ACGCGCCAGAAAGGCGCAAGAGCTGGCGGTGTCACCTGCGCCGAGCTGCTGCTGCTG	1043		
Db	541	ACGCGCCAGAAAGGCGCAAGAGCTGGCGGTGTCACCTGCGCCGAGCTGCTGCTGCTG	600		
QY	1044	CGCGACCGGGGCTTGGCGCGCGCGAGCGCGCCGAGTTGAGCGGCTCTGGGTGGCCTTC	1103		



Db	601	CGCGACCGGGGCTGCGCCGCGACGAGCGCGCGAGTTGACGCGCTTGCGGTGCCTTC	660
OY	1104	TGCGGCTGCTGGAACCTGCTGAGACGAGCATATGCGACGCTGCTGGAAGCTGGCGCGCG	1163
-Db	661	TGCGGCTGCTGGAACCTGCTGAGACGAGCATATGCGACGCGCTGGAAGCTGGCGCGCG	720
OY	1164	TTTCCGCTGCAAGCGCGCGCGCGAGCCGCTGGTGGCAAGAGTGTGGTGGCGCTTC	1223
Db	721	TTTCCGCTGCAAGCGCGCGCGCGAGCCGCTGGTGGCAAGAGTGTGGTGGCGCTTC	780
OY	1224	GCGGTGGGCGCGCGCGCTGAGGACCCCGACGCTGGCGCTGCAGGCGGAGGCGCATTC	1283
Db	781	GCGGTGGGCGCGCGCGCTGAGGACCCCGACGCTGGCGCTGCAGGCGGAGGCGCATTC	840
OY	1284	GACGTGCGGAGCTGCGGAGCTGAGAGCGCGAGTCTTTCAGTGTGGGCGAGATGATCGAC	1343
Db	841	GACGTGCGGAGCTGCGGAGCTGAGAGCGCGAGTCTTTCAGTGTGGGCGAGATGATCGAC	900
OY	1344	AACATGAGATGAGAGTCAACGTGCCCCGCTGAGCCGTGCAAGCCCGGCAAGCGCGCGG	1403
Db	901	AACATGAGATGAGAGTCAACGTGCCCCGCTGAGCCGTGCAAGCCCGGCAAGCGCGCGG	960
OY	1404	GCGGAGCTTCCTGTCCAAGGTCAAGCGCGCGCGCTCTCGGTGCTGTCTTTCAGAGAGCG	1463
Db	961	GCGGAGCTTCCTGTCCAAGGTCAAGCGCGCGCGCTCTCGGTGCTGTCTTTCAGAGAGCG	1020
OY	1464	GCGGAGGAGTGTGAGACCCCGAGAGGACCTGGCGCGCATCTTTTCGCGCGGTGCTGTG	1523
Db	1021	GCGGAGGAGTGTGAGACCCCGAGAGGACCTGGCGCGCATCTTTTCGCGCGGTGCTGTG	1080
OY	1524	GCGGCTGTGACCTAGACCGTGTGCGTGGCGAGCTGAGCTGACAGACACCGACCGCGC	1583
Db	1081	GCGGCTGTGACCTAGACCGTGTGCGTGGCGAGCTGAGCTGAGACAGACACCGACCGCGC	1140
OY	1584	CTGCTGCTGCGCGCTCTCTCTCCCTGAGAAAAGACTTGGGATGGGTGTGGGCTGTGCTGT	1643
Db	1141	CTGCTGCTGCGCGCTCTCTCTCCCTGAGAAAAGACTTGGGATGGGTGTGGGCTGTGCTGT	1200
OY	1644	GCGAGGGGAGTGTCTTAAACCCCGGTGTGACATGGGTGACACGCGCGTTTCAGTGCAC	1703
Db	1201	GCGAGGGGAGTGTCTTAAACCCCGGTGTGACATGGGTGACACGCGCGTTTCAGTGCAC	1260
-OY	1704	ATCTGCTGAGGACGAGACAGGTTTCTCTTGCTGGCCCGGAGAGATTAACTTTGCGCC	1763
Db	1261	ATCTGCTGAGGACGAGACAGG-TTCTCTTGCTGGCCCGGAGAGATTAACTTTGCGCC	1319
OY	1764	GCGCGTCAAGGCAATACGCTTAACGTCTGACAGACTTTATTCCTAATTAATGAANAAC	1823
Db	1320	GCGCGTCAAGGCAATACGCTTAACGTCTGACAGACTTTATTCCTAATTAATGAANAAC	1379
OY	1824	GTCAAGAGGACCTCTAGATCCCTCGAGTTAAATGATTTAAACAATGTCTGTGGGCGCTC	1883
Db	1380	GTCAAGAGGACCTCTAGATCCCTCGAGTTAAATGATTTAAACAATGTCTGTGGGCGCTC	1439
OY	1884	TTTACAGGGAGTCCGAGTTCCGTTGCCACCCCTGCGACGTCGCGCCCTTTCTGCGTGGG	1943
Db	1440	TTTACAGGGAGTCCGAGTTCCGTTGCCACCCCTGCGACGTCGCGCCCTTTCTGCGTGGG	1499
OY	1944	ACAGTTTGAAGGTTGGGTGGGTGAGTGAAGTTTGAAGAGGACGCTGTTGGTTCTA	2003
Db	1500	ACAGTTTGAAGGTTGGGTGGGTGAGTGAAGTTTGAAGAGGACGCTGTTGGTTCTA	1559
OY	2004	TGTGGTGTCTGTTCCTCCGCGACAAGAAAAATGGAATCAATGTCAACACTTTTATTA	2063
Db	1560	TGTGGTGTCTGTTCCTCCGCGACAAGAAAAATGGAATCAATGTCAACACTTTTATTA	1619
OY	2064	CCTTAATCTTTACAGGCGCTAAATTTAGAGAGTGTCTGAGAGCACTTATACAAAGGCG	2123
Db	1620	CCTTAATCTTTACAGGCGCTAAATTTAGAGAGTGTCTGAGAGCACTTATACAAAGGCG	1679
OY	2124	TTTCTTAAGACGCGCTAACGCTTCTTCAAGAGATTATTCATTCGTCCCAAGAGCA	2183

Db	1680	TTCTCTTAAGACGGGCTACAGCCCTTCCTAGCAGAGTTTATTCATTCTGTCGCCAAGAGCA	1732
Qy	2184	GCTAGAAAGATTTTGAGGTCATGACTCTCCCACTGCGGCTCAAGGGGCTGACCTTATTTAGG	2243
Db	1740	GCTAGAAAGATTTTGAGGTCATGACTCTCCCACTGCGGCTCAAGGGGCTGACCTTATTTAGG	1799
Qy	2244	AAACCAAAAGAGGGGTGGTTTAAACCTACTCTCAACGAAATTGAGATCAGTGGGCACACTTGC	2303
Db	1800	AAACCAAAAGAGGGGTGGTTTAAACCTACTCTCAACGAAATTGAGATCAGTGGGCACACTTGC	1859
Qy	2304	CTGGCGAAAAGGGGCTCTCCCAAGCCACCCGAGATGGGGGGTAAAGAGAAAGCAGAGGCT	2363
Db	1860	CTGGCGAAAAGGGGCTCTCCCAAGCCACCCGAGATGGGGGGTAAAGAGAAAGCAGAGGCT	1919
Qy	2364	TGGGGTAAAGGGCCACTGGTGTGTTTAAACAGGCACTTCTCTCTCTGAGGGCTTATTTTTC	2423
Db	1920	TGGGGTAAAGGGCCACTGGTGTGTTTAAACAGGCACTTCTCTCTCTCTGAGGGCTTATTTTTC	1979
Qy	2424	TTCAAGAACTAGACCAAGAGTGTGTTGAACCTCTTTTGACAGAGGGGTGGGAATCCTCTTTAG	2483
Db	1980	TTCAAGAACTAGACCAAGAGTGTGTTGAACCTCTTTTGACAGAGGGGTGGGAATCCTCTTTAG	2039
Qy	2484	AGCACTTAATCTTAATTTATCCCGGAATGAGGTGCTGACAGATAGAGGGGCTGCTTT	2543
Db	2040	AGCACTTAATCTTAATTTATCCCGGAATGAGGTGCTGACAGATAGAGGGGCTGCTTT	2099
Qy	2544	GGCAGCTCCCTGACCCCGCGGCTGCGCCCTCTCCGGGGTAAATGTGGCACTTACTGCGCCA	2603
Db	2100	GGCAGCTCCCTGACCCCGCGGCTGCGCCCTCTCCGGGGTAAATGTGGCACTTACTGCGCCA	2159
Qy	2604	CAGAGGTTTTGAGCCAAATCAGCTCTGACCTGGGTTGAATGTAAACAGCTTTAACTTGGG	2663
Db	2160	CAGAGGTTTTGAGCCAAATCAGCTCTGACCTGGGTTGAATGTAAACAGCTTTAACTTGGG	2219
Qy	2664	ATTTAAGAAGCTTTTAAAGGTAATATCTCTGAAAGAAAATAGAGTAACCAACAGCT	2723
Db	2220	ATTTAAGAAGCTTTTAAAGGTAATATCTCTGAAAGAAAATAGAGTAACCAACAGCT	2279
Qy	2724	GTACTATGAAGAGCTGTATTATTTTATAAAGAACGCTGGGCCATGAACTC	2771
Db	2280	GTACTATGAAGAGCTGTATTATTTTATAAAGAACGCTGGGCCATGAACTC	2327
RESULT 7			
AK127646		2327 bp	MRNA
LOCUS			linear
DEFINITION	Homo sapiens cDNA FLJ545744 f18, clone KIDNE2017153.		PRI 19-FEB-2004
ACCESSION	AK127646		
VERSION	AK127646.1	GI:34534649	
KEYWORDS	oligo capping; f18 (full insert sequence).		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Bukaryovc, Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1		
	Oshima, A., Takahashi, Fujii, A., Tanase, T., Imose, N., Takeuchi, K., Arita, M., Mueashino, K., Yuuki, H., Hara, H., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Nakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai, Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Watanabe, M., Murakawa, K., Kanehori, K., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Maehuo, Y., Nagai, K. and Iwagui, T.		
TITLE	NEBO human cDNA sequencing project		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 2327)		
AUTHORS	Iwagui, T. and Yamamoto, J.		
TITLE	Direct Submissions		
JOURNAL	Submitted (15-JUL-2003) Takao Iwagui, FLJ Project (HRI Team); 2-6-7		
COMMENT	Kazusa-Kamatari, Katsarazu, Chiba 232-0818, Japan (E-mail:genomdb@hri.co.jp, Tel:81-438-52-3975, Fax:81-438-52-3986)		
	NEBO human cDNA sequencing project supported by Ministry of		



Economy, Trade and Industry of Japan: cDNA full insert sequencing:  
Research Association for Biotechnology (RAB); cDNA library  
construction: Helix Research Institute (HRI) (supported by Japan  
Key Technology Center etc.); 5' - & 3'-end one pass sequencing: RAB,  
HRI, and Biotechnology Center, National Institute of Technology and  
Evaluation; clone selection for full insert sequencing: HRI and  
RAB; annotation: Reverse Proteomics Research Institute, HRI and  
RAB.

FEATURES  
source

CDS

Location/Qualifiers  
1..2327  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="KIDNE2017153"  
/issue\_type="Kidney"  
/clone\_1ib="KIDNE2"  
/note="Cloning vector: pME18SFL3"  
415..1122  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="BAC87072.1"  
/db\_xref="GI:34534650"  
/translation="MARBECKALDGLNKTTACTYHHLVLTWGSADSQNTLOEOKTR  
OKAELAVSTCARLTAVLRDRGLADBEAFERLWVAFSGCIDLEADMRRLAETGAA  
PLHAPRRLVETGVAGASGVAFALSTRSLRLEAGDFVADLRLEIREVLYOVEM  
IDNEMKVNVPWTYQARQAAAEILSTVSAPSSVSLQERGGCDPRKALALFLG  
AVLAAVALAVCAKLS"

## ORIGIN

Query Match 62.5%; Score 1952; DB 8; Length 2327;  
Best Local Similarity 99.7%; Pred. No. 0;

Matches 2322; Conservative 0; Mismatches 5; Indels 1; Gaps 1;

444 ACCCGGCGCCCTTGGACGCGCTTAAGCGAGCGCGGCTCTGACGCTTGTCCCC 503  
1 ACCCGGCGCCCTTGGACGCGCTTAAGCGAGCGCGGCTCTGACGCTTGTCCCC 60  
504 GGAAGTTGGACCCACGAGGATGAGGACCGGACCTCGAGCTTGGAGGAGCACTGG 563  
61 GGAAGTTGGACCCACGAGGATGAGGACCGGACCTCGAGCTTGGAGGAGCACTGG 120  
564 AGGCGAGGCGGATGACAGACAGAGTGTGACTGAGTGTGCGCTTGGAGGAGTGAAG 623  
121 AGGCGAGGCGGATGACAGACAGAGTGTGACTGAGTGTGCGCTTGGAGGAGTGAAG 180  
624 AGGAGCGGCGGATGACAGAGTGTGACTGAGTGTGCGCTTGGAGGAGTGAAGT 683  
181 AGGAGCGGCGGATGACAGAGTGTGACTGAGTGTGCGCTTGGAGGAGTGAAGT 240  
684 CGAGGGGTCCGAGCGGCGCTCCGTGACGTTGGCGGTGAGCGAGGAGTGAAGTCA 743  
241 CGAGGGGTCCGAGCGGCGCTCCGTGACGTTGGCGGTGAGCGAGGAGTGAAGTCA 300  
744 GAGAGCGTTCGTGCGCGCGGCCCAAGCGCGGATGAGGAGTGAAGTCAAGTCCGCG 803  
301 GAGAGCGTTCGTGCGCGCGGCCCAAGCGCGGATGAGGAGTGAAGTCAAGTCCGCG 360  
804 CTGAGGGGAGGCTTAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 863  
361 CTGAGGGGAGGCTTAACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 420  
864 AGGAGGAGTGAAGGCGCGCTGCTGACGCGGCTCAACAAAGACAGCTGCGGTACCA 923  
421 AGGAGGAGTGAAGGCGCGCTGCTGACGCGGCTCAACAAAGACAGCTGCGGTACCA 480  
924 CTGAGTCTGACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 983  
481 CTGAGTCTGACCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
984 AGCGCGCAGAGGCGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1043  
541 AGCGCGCAGAGGCGGAGGAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600

QY 1044 GCGACCGGCGGCGCTGCGCGCGCGAGAGGCGCGAGTTCGAGCGGCTTGGGTCCTTC 1103  
DB 601 GCGACCGGCGGCGCTGCGCGCGCGAGAGGCGCGAGTTCGAGCGGCTTGGGTCCTTC 660  
QY 1104 TCGGCTGCTGCGAGCTTGTCTGAGGAGCGAGCACTGCGCTTGTGAGTGTGCGCG 1163  
DB 661 TCGGCTGCTGCGAGCTTGTCTGAGGAGCGAGCACTGCGCTTGTGAGTGTGCGCG 720  
QY 1164 TTCCGCTGCGACCGGCGCGCGCGAGCCGCTGCTGCGACAGTGTGCTGCGCGCTTC 1223  
DB 721 TTCCGCTGCGACCGGCGCGCGCGAGCCGCTGCTGCGACAGTGTGCTGCGCGCTTC 780  
QY 1224 GCGCGTGGCGGCGCGCGCGCTGAGACCCGCGAGCTTGTGAGGCGGAGGCGACTTC 1283  
DB 781 GCGCGTGGCGGCGCGCGCGCTGAGACCCGCGAGCTTGTGAGGCGGAGGCGACTTC 840  
QY 1284 GACGTGCGGAGCTTGTGCGGAGCTGAGAGCGAGGCTCTTCAAGTGTGAGGAGTGA 1343  
DB 841 GACGTGCGGAGCTTGTGCGGAGCTGAGAGCGAGGCTCTTCAAGTGTGAGGAGTGA 900  
QY 1344 AACATGAGATGAAGATGAAGTCAAGTCCCGCTGCGACCGTGTCAAGCCCGCGAG 1403  
DB 901 AACATGAGATGAAGATGAAGTCAAGTCCCGCTGCGACCGTGTCAAGCCCGCGAG 960  
QY 1404 GCGAGGCTCTGTGCGACGAGTCAAGCGCGCGCGCTCTGCTGCTGCTGCTGCTGCT 1463  
DB 961 GCGAGGCTCTGTGCGACGAGTCAAGCGCGCGCGCTCTGCTGCTGCTGCTGCTGCT 1020  
QY 1464 GCGGCGGCTTGTGCGACCCCGAGAGGCGCTGCGCGCATCTTTTGGCGCGCTGCTG 1523  
DB 1021 GCGGCGGCTTGTGCGACCCCGAGAGGCGCTGCGCGCATCTTTTGGCGCGCTGCTG 1080  
QY 1524 GCGGCTGTGCGCTTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1583  
DB 1081 GCGGCTGTGCGCTTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1140  
QY 1584 CTGCTGCTGCGCGCTTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1643  
DB 1141 CTGCTGCTGCGCGCTTGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1200  
QY 1644 GCAAGGAGGATGTCTTAAACCCCGTGTGTGATGAGTCAAGCGGCTTTCAGTGCAC 1703  
DB 1201 GCAAGGAGGATGTCTTAAACCCCGTGTGTGATGAGTCAAGCGGCTTTCAGTGCAC 1260  
QY 1704 ATTCGCTGCGGCGAGGACAGCGGCTTGTGCGCGCGCGCGCGCGCGCGCGCGCG 1763  
DB 1261 ATTCGCTGCGGCGAGGACAGG-ATTCGCTGCGGCGCGCGCGCGCGCGCGCGCG 1319  
QY 1764 GCGCGTCAAGGCACTTACCGCTAACGCTGACAGAGCTTATTCCTATTAATGAGAAC 1823  
DB 1320 GCGCGTCAAGGCACTTACCGCTAACGCTGACAGAGCTTATTCCTATTAATGAGAAC 1379  
QY 1824 GTCAAGTGAACCTTAAATCCCTCGAGTTAATGATTAAACATGTGCTGTTGGGCGTC 1883  
DB 1380 GTCAAGTGAACCTTAAATCCCTCGAGTTAATGATTAAACATGTGCTGTTGGGCGTC 1439  
QY 1884 TTTAACAGGAGTCCGAGTTCGCTGCGCACCCCTGCGAGCGTCCGCTTCTGCGTGG 1943  
DB 1440 TTTAACAGGAGTCCGAGTTCGCTGCGCACCCCTGCGAGCGTCCGCTTCTGCGTGG 1499  
QY 1944 ACAATTTGAAAAAGTGTGAGGAGTGAAGTGAAGTGAAGGAGCGCTGTTGGTCTTA 2003  
DB 1500 ACAATTTGAAAAAGTGTGAGGAGTGAAGTGAAGTGAAGGAGCGCTGTTGGTCTTA 1559  
QY 2004 TGTGTTGTCTGTTTCCCGAGCAAGAAAAATTCATCAATGTCAGAGCTTTATTA 2063  
DB 1560 TGTGTTGTCTGTTTCCCGAGCAAGAAAAATTCATCAATGTCAGAGCTTTATTA 1619  
QY 2064 CTTTAATCTTTCAAGGCGCTTAATTTAAGAGTGTCTGAGAGGAGTTCATCAAGAGGC 2123  
DB 1620 CTTTAATCTTTCAAGGCGCTTAATTTAAGAGTGTCTGAGAGGAGTTCATCAAGAGGC 1679  
QY 2124 TTTCTCTAAGAGCGGCTTACAGCCCTTCTTACAGAGTTATTCATTCGTCCTCAAGACA 2183

Db	1680	TTTCTTAAGACGCGCTAAGCCCTTCTTAGCAAGCTTTATCATTGTCCTCCCAAGACA	17398
Qy	2184	GCTAGAAAGATTGAGGTCAATGACCTCCCACTGCGCTCAGGGCTGACCTTATTTAG	2243
Db	1740	GCTAGAAAGATTGAGGTCAATGACCTCCCACTGCGCTCAGGGCTGACCTTATTTAG	17998
Qy	2244	AAACCAAGAGGGGTGGGTTGAACCTACTCTCAGCGACTTTGGATCCAGTGGCACA	2303
Db	1800	AAACCAAGAGGGGTGGGTTGAACCTACTCTCAGCGACTTTGGATCCAGTGGCACA	1859
Qy	2304	CTGGCGAAAAGGGGCTCTCCCGACGACCCGGAGATGGGGGTAAAGAAAGCAGAGGCT	2363
Db	1860	CTGGCGAAAAGGGGCTCTCCCGACGACCCGGAGATGGGGGTAAAGAAAGCAGAGGCT	1919
Qy	2364	TGGGGTAAAGGCGCACTGGGTGTTTAAACAGGCACTTCTCTCTGCGGCTTATTTTG	2423
Db	1920	TGGGGTAAAGGCGCACTGGGTGTTTAAACAGGCACTTCTCTCTGCGGCTTATTTTG	1979
Qy	2424	TTCAGAACTTAAACACAGAGTGTTTGAACCTCTTTGACAGAGGCTGGAAATCTCTTTAG	2483
Db	1980	TTCAGAACTTAAACACAGAGTGTTTGAACCTCTTTGACAGAGGCTGGAAATCTCTTTAG	2039
Qy	2484	AGCACTTAAATCCTATTTATATCCCTGGAAATGTCGGTGGTGGCAGTAAAGGCTGGCTTT	2543
Db	2040	AGCACTTAAATCCTATTTATATCCCTGGAAATGTCGGTGGTGGCAGTAAAGGCTGGCTTT	2099
Qy	2544	GGCAGCTCCCTGACCCCGCGCTGGCGCCCTCCGGGGTAAATGTCGACTTACTGGCCA	2603
Db	2100	GGCAGCTCCCTGACCCCGCGCTGGCGCCCTCCGGGGTAAATGTCGACTTACTGGCCA	2155
Qy	2604	CAGAGGTTTGAAGCCAAATCAGCTCTGAGACTGGGTTAGAAATGTAAACAGCTTTAACTTGGG	2663
Db	2160	CAGAGGTTTGAAGCCAAATCAGCTCTGAGACTGGGTTAGAAATGTAAACAGCTTTAACTTGGG	2219
Qy	2664	ATTTAAGAAAGCTTTTAAAGGTAAATATATCTCTGAAAAGAAATGACGTAAACACAGGCT	2723
Db	2220	ATTTAAGAAAGCTTTTAAAGGTAAATATATCTCTGAAAAGAAATGACGTAAACACAGGCT	2279
Qy	2724	GTACTATGAAGAGCTGTATTTTAAATTAAGAAAGCTGGGCAATGAAGCTC 2771	
Db	2280	GTACTATGAAGAGCTGTATTTTAAATTAAGAAAGCTGGGCAATGAAGCTC 2327	
RESULT 8			
CQ748891			
LOCUS	CQ748891	708 bp	DNA
DEFINITION	Sequence 34825 from Patent WO02068579.	linear	PAT 03-PBB-2004
ACCSSION	CQ748891		
VERSION	CQ748891.1	GI:42376058	
KEYWORDS			
SOURCE			
ORGANISM	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Homidae; Homo.		
REFERENCE	1		
AUTHORS	Venter, C.J., Adams, M.C., Li, P.W. and Myers, R.W.		
TITLE	Ki68, such as nucleic acid arrays, comprising a majority of		
	humanexons or transcripts, for detecting expression and other uses		
	thereof		
JOURNAL	Patent: WO 02068579-A 34825 06-SEP-2002;		
PE Corporation (NY) (US)			
FEATURES	location/Qualifiers		
source	1..708		
	/organism="Homo sapiens"		
	/mol_type="unassigned DNA"		
	/db_xref="taxon:9606"		
ORIGIN			
Query Match	21.0%;	Score 657;	DB 6; Length 708;
Best Local Similarity	99.9%;	Pred. No. 0;	
Matches 707;	Conservative	0;	Mismatches 1; Indels 0; Gaps 0;

QY	858	ATGCGAGAGGAGAGAGTGCAGAGGCGCTGCTGAGACGGGCTCAACAGAGAGACTGGCTGCTAC	917
Db	1	ATGCGAGAGGAGAGAGTGCAGAGGCGCTTCTGAGCGGGCTCAACAGAGAGACTGGCTGCTAC	60
QY	918	CACCAACCTGGGCTGAGACCGTCGAGTGGGCTCGGCGGAGCTCGAGAACTCGGCGAGAGACTGG	977
Db	61	CACCACTTGAGTGTACTGACCGTCGAGTGGGCTCGGCGGAGCTCGAGAACTCGGCGAGAGACTGG	120
QY	978	CAAAAGACGCGCCAGAAAGCGCAGAGAGCTGGCGGTGTCCACTTGGCGCCGGCTGACTGCT	1037
Db	121	CAAAAGACGCGCCAGAAAGCGCAGAGAGCTGGCGGTGTCCACTTGGCGCCGGCTGACTGCT	180
QY	1038	GTGCTGGGCGGACCGGGGCGCTGGGCGCTGAGCGAGCGCGCCGAGTTTGAACCGGCTTGGGCTG	1097
Db	181	GTGCTGGGCGGACCGGGGCGCTGGGCGCTGAGCGAGCGCGCCGAGTTTGAACCGGCTTGGGCTG	240
QY	1098	GCTTTCTCGGGCTGCTGAGACTGTGTGAGAAAGCGGACATGGGAGCGCTGGCTGAGACTGGGCG	1157
Db	241	GCTTTCTCGGGCTGCTGAGACTGTGTGAGAAAGCGGACATGGGAGCGCTGGCTGAGACTGGGCG	300
QY	1158	GCCGCGTGTCCCGCTGACACGCGCCGCGGCGACCGCTGTGTGGGCAACAGGTGTGTGGCTGGCGCC	1217
Db	301	GCCGCGTGTCCCGCTGACACGCGCCGCGGCGCGCTGTGTGGGCAACAGGTGTGTGGCTGGCGCC	360
QY	1218	TCTTTCGGGCGTGGCGGCGCGGCGCTGTAGACACCCGACAGCTTGGCGCTTCGAGGCGAGAGCG	1277
Db	361	TCTTTCGGGCGTGGCGGCGCGGCGCGCTGTAGACACCCGACAGCTTGGCGCTTCGAGGCGAGAGCG	420
QY	1278	GACTTTCGACGCTGCGGAGACTGTGCGGGGACCTGGAAGGCGAGAGTCTTTCAGAGTGGGCGGAGATG	1337
Db	421	GACTTTCGACGCTGCGGAGACTGTGCGGGGACCTGGAAGGCGAGAGTCTTTCAGAGTGGGCGGAGATG	480
QY	1338	ATCGACACACATGAGAGATGAGAGGTCAACGTGCGCCGCTGACACCGTGCAGAGCCCGGACAGCG	1397
Db	481	ATCGACACACATGAGAGATGAGAGGTCAACGTGCGCCGCTGACACCGTGCAGAGCCCGGACAGCG	540
QY	1398	GCGGCGCGCGGAGCTCTGTGTCCAGCGTCAAGCGCGCGGCGCTTCTTCGATCGTGTCTTGGCAG	1457
Db	541	GCGGCGCGCGGAGCTCTGTGTCCAGCGTCAAGCGCGCGGCGCTTCTTCGATCGTGTCTTGGCAG	600
QY	1458	GAGGCGCGGCGGGGGGTTTGCACGCCCGAGAGAGCCCTGGCGCGCGCAATCTTTTCGGGCGCGTG	1517
Db	601	GAGGCGCGGCGGGGGGTTTGCACGCCCGAGAGAGCCCTGGCGCGCGCAATCTTTTCGGGCGCGTG	660
QY	1518	CTGCTGAGCGGCTGTGAGCCCTTAGCCGTTGCGTGGCGAGCTGAGCTGA	1565
Db	661	CTGCTGAGCGGCTGTGAGCCCTTAGCCGTTGCGTGGCGAGACTGAGCTGA	708
RESULT 9			
HMMZ85E06			
LOCUS	HMMZ85E06	655 bp	mRNA linear PRI 29-AUG-1998
DEFINITION	Homo sapiens full length insert cDNA clone ZB85E06.		
VERSION	AF086461		
KEYWORDS	AF086461.1 GI:3483806		
ORGANISM	FLI CDNA.		
	Homo sapiens (human)		
	Homo sapiens		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;		
	Hominae; Homo.		
	1 (bases 1 to 655)		
REFERENCE	Mossner, J., Tan, F., Marra, M., Kucaba, T., Yandell, M., Martin, J.,		
AUTHORS	Math, G., Bowles, L., Wylie, T., Bowers, Y., Steptoe, M., Theising, B.,		
	Geisel, S., Allen, M., Underwood, K., Chappell, V., Person, B.,		
	Gibbons, M., Harvey, N., Pape, D., Chamberlain, A., Morales, R.,		
	Schub, R., Rletter, E., Kohn, S., Swaller, T., Behnmer, K., Hillier, L.,		
	Wilson, R. and Waterston, R.		
TITLE	Full Clone Sequencing of the Longest Available Member from Each		
JOURNAL	Unigene Cluster		
REFERENCE	2 (bases 1 to 655)		

AUTHORS Waterston, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (24-06-1998) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 COMMENT SUBMITTED BY: Genome Sequencing Center  
 Department of Genetics  
 Washington University  
 St. Louis MO 63108, USA  
 http://genome.wustl.edu/gsc  
 mailto:est@wustl.wustl.edu

NOTICE: This sequence represents the full insert of this cDNA. No attempt has been made to verify whether this corresponds to the full-length of the original mRNA from which it was derived. We have tried to obtain double-stranded, or double chemistry sequence across the entire clone, but potentially, there are areas in the sequence where this level of coverage was not achieved. Nevertheless, we are confident of the accuracy of this sequence as all regions of low quality, as defined by PHRAP (P. Green, in preparation), were visually inspected and edited accordingly. The consensus quality values for this sequence have been submitted separately.

FEATURES  
 The location of this clone is unknown.  
 location/Qualifiers

1..655  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="TMAB:347458"  
 /clone\_1lb="Soares\_fetal\_heart\_NbHH19W"

## ORIGIN

Query Match 20.7%; Score 646; DB 8; Length 655;  
 Best Local Similarity 100.0%; Pred. No. 0;

Matches 646; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2129 CTAAGACGGCTACAGCCCTTCTCTAGCAGAGTTTATCCATTGTCCTCCCAAGAGCACTG 2188  
 DB 1 CTAAGACGGCTACAGCCCTTCTCTAGCAGAGTTTATCCATTGTCCTCCCAAGAGCACTG 60  
 QY 2189 AAGAGATTGAGGTATGACCTCCACCTGCGCTGAGGGGCTGACCCCTATTAGAAAC 2248  
 DB 61 AAGAGATTGAGGTATGACCTCCACCTGCGCTGAGGGGCTGACCCCTATTAGAAAC 120  
 QY 2249 AAAAGAGGTGGTTGAACCTACTCTGACGCACTTGATCAGTGGCACACTTGCTGCG 2308  
 DB 121 AAAAGAGGTGGTTGAACCTACTCTGACGCACTTGATCAGTGGCACACTTGCTGCG 180  
 QY 2309 GAAAAAGGCTCTCCCAAGCACCAGAGATGGGGGTTAAGAGAGAGAGAGAGGCTTGGG 2368  
 DB 181 GAAAAAGGCTCTCCCAAGCACCAGAGATGGGGGTTAAGAGAGAGAGAGAGGCTTGGG 240  
 QY 2369 TAGGGCACTGCTGTTAAACAGGCACTTCTCTCTGCGGCTTATTTTGTTCAG 2428  
 DB 241 TAGGGCACTGCTGTTAAACAGGCACTTCTCTCTGCGGCTTATTTTGTTCAG 300  
 QY 2429 AACTAGACAGAGTGTGAACCTCTCTTTCAGAGAGGGCTGGGAATCTCTTTAGAGCAC 2488  
 DB 301 AACTAGACAGAGTGTGAACCTCTCTTTCAGAGAGGGCTGGGAATCTCTTTAGAGCAC 360  
 QY 2489 TTAATCTTATTTATCCCTGGAATGAGCTGCGCAGTAGAGGGCTGGCTTGGCAG 2548  
 DB 361 TTAATCTTATTTATCCCTGGAATGAGCTGCGCAGTAGAGGGCTGGCTTGGCAG 420  
 QY 2549 CTCCTGACCCCGCGCTGCGCCCTCCGCGGTATATGGCACTTACGCGCCACAGAG 2608  
 DB 421 CTCCTGACCCCGCGCTGCGCCCTCCGCGGTATATGGCACTTACGCGCCACAGAG 480  
 QY 2609 GTTTTGGCAATCAGCTCTGAGACTGGGTTAGATGTAACAGCTTTAACTTGGATT 2668  
 DB 481 GTTTTGGCAATCAGCTCTGAGACTGGGTTAGATGTAACAGCTTTAACTTGGATT 540

QY 2669 AGAAGCTTTAAAGCTAATAATCTCTGAAAGAAAAATGACGTAAACAGACCGGTACT 2728  
 DB 541 AGAAGCTTTAAAGCTAATAATCTCTGAAAGAAAAATGACGTAAACAGACCGGTACT 600  
 QY 2729 ATGAAGCTGTATTATTTAATAAGAACGCTGGGCCATGACTCAT 2774  
 DB 601 ATGAAGCTGTATTATTTAATAAGAACGCTGGGCCATGACTCAT 646

RESULT 10  
 LOCUS CO746021 708 bp DNA linear PAT 03-FEB-2004  
 DEFINITION Sequence 31955 from Patent WO02068579.  
 ACCESSION CO746021  
 VERSION CO746021.1 GI:42365754  
 KEYWORDS  
 SOURCE  
 ORGANISM Homo sapiens (human)

REFERENCE  
 AUTHORS Ventner, C.J., Adams, M.C., Li, P.W. and Myers, E.W.  
 TITLE Kites, such as nucleic acid arrays, comprising a majority of humanexons or transcripts, for detecting expression and other uses thereof  
 Patent: WO 02068579-A 31955 06-SEP-2002;  
 PE Corporation (NY) (US)

FEATURES  
 source  
 location/Qualifiers  
 1..708  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"

## ORIGIN

Query Match 19.4%; Score 606; DB 6; Length 708;  
 Best Local Similarity 99.7%; Pred. No. 0;

Matches 706; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 858 ATGGGAGAGAGAGAGTGAAGAGCGCTGCTGAGCGGCTCAACAAAGACATGCGTGTAC 917  
 DB 1 ATGGGAGAGAGAGAGTGAAGAGCGCTGCTGAGCGGCTCAACAAAGACATGCGTGTAC 60  
 QY 918 CACGACCTGCTGCTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 977  
 DB 61 CACGACCTGCTGCTGACCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 120  
 QY 978 CAAAAGAGCGCCAGAAAGCGCAGAGAGCTGAGCGGTGTCAACTGCGCCCGCTGACTGCT 1037  
 DB 121 CAAAAGAGCGCCAGAAAGCGCAGAGAGCTGAGCGGTGTCAACTGCGCCCGCTGACTGCT 180  
 QY 1038 GTGCTGCGGACCGGGGCTTGGCGCGCAGACGCGCGCGAGTTGAGCGGCTTGGGTG 1097  
 DB 181 GTGCTGCGGACCGGGGCTTGGCGCGCAGACGCGCGCGAGTTGAGCGGCTTGGGTG 240  
 QY 1098 GCTTCTGCGGCTGCTGAGACCTGCTGAGAGGAGCATGCGACGCTGCTGAGAGCTGAGC 1157  
 DB 241 GCTTCTGCGGCTGCTGAGACCTGCTGAGAGGAGCATGCGACGCTGCTGAGAGCTGAGC 300  
 QY 1158 GCGCGGTCGCGCTGCAAGCGCGCGCGGCGAGCGCTGCTGAGAGCTGCTGAGAGCTGAGC 1217  
 DB 301 GCGCGGTCGCGCTGCAAGCGCGCGCGGCGAGCGCTGCTGAGAGCTGCTGAGAGCTGAGC 360  
 QY 1218 TCTTCCGCGTGGCGCGCGCGCGCTGAGCACTGCGAGCTGCGGCTGAGAGCGAGAGG 1277  
 DB 361 TCTTCCGCGTGGCGCGCGCGCGCTGAGCACTGCGAGCTGCGGCTGAGAGCGAGAGG 420  
 QY 1278 GACTTGGAGCTGCGGAGCTGCGGAGGCTGAGAGCGCGAGGCTTCAAGTGGGCGAGATG 1337  
 DB 421 GACTTGGAGCTGCGGAGCTGCGGAGGCTGAGAGCGCGAGGCTTCAAGTGGGCGAGATG 480  
 QY 1338 ATCGAACATGAGAGTGAAGGTCAACGTGCCCGCTGGAACCGTGGCAAGCCCGAGGCG 1397

Db 481 ATCGACATGAGATGAAGGTCAACGTCGCCCGCTGAGACCTGACCCGGCAGGCG 540  
Qy 1398 GGGGGGCGGAGCTCTCTGTCACAGGTCAGCGCGCGCCCTCTCGTCTGTCCTTGCG 1457  
Db 541 GGGGGGCGGAGCTCTGTCACAGGTCAGCGCGCGCCCTCTCGTCTGTCCTTGCG 600  
Qy 1458 GAGCGCGGGGGGTTGCGACACCGGAGAGCGCTTGCGCGCCATCTTTTGCGCGCGT 1517  
Db 601 GAGCGCGGGGGGTTGCGACACCGGAGAGCGCTTGCGCGCCATCTTTTGCGCGCGT 660  
Qy 1518 CTGCTGCGGCTGTGTGCGCCCTAGCCGTGTGCGCGACCTGAGCTGA 1565  
Db 661 CTGCTGCGGCTGTGTGCGCCCTAGCCGTGTGCGCGACCTGAGCTGA 708  
RESULT 11  
LOCUS AC027340 216441 bp DNA linear HTG 15-JUL-2000  
DEFINITION Homo sapiens chromosome 19 clone CTD-2257C19, WORKING DRAFT  
AC027340  
AC027340.2 GI:9211228  
VERSION HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 216441)  
AUTHORS DOE Joint Genome Institute.  
TITLE Sequencing of Human Chromosome 19  
JOURNAL Unpublished  
AUTHORS DOE Joint Genome Institute.  
REFERENCE Direct Submission  
AUTHORS Submitted (30-MAR-2000) Production Sequencing Facility, DOE Joint  
TITLE Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94596, USA  
JOURNAL On Jul 15, 2000 this sequence version replaced gi:7341654.  
COMMENT -----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: <http://www.jgi.doe.gov>  
-----  
Project Information  
Center Project Name: 726304, BC691328  
Center clone name: CITB-HI\_2257C19  
-----  
Summary Statistics  
Consensus quality: 164680 bases at least Q40  
Consensus quality: 187929 bases at least Q30  
Consensus quality: 195340 bases at least Q20  
Estimated insert size: 211341; sum-of-contigs estimation  
Quality coverage: 8.59 in Q20 bases; agarose-1p estimation  
Quality coverage: 6.46 in Q20 bases; sum-of-contigs estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 52 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 1041: contig of 1041 bp in length  
\* 1042 1141: gap of unknown length  
\* 1142 2370: contig of 1229 bp in length  
\* 2371 2470: gap of unknown length  
\* 2471 3552: contig of 1082 bp in length  
\* 3553 3653: gap of unknown length  
\* 3653 4683: contig of 1031 bp in length  
\* 4684 4783: gap of unknown length  
\* 4784 6178: contig of 1395 bp in length

\* 6179 6278: gap of unknown length  
\* 6279 7442: contig of 1164 bp in length  
\* 7443 7542: gap of unknown length  
\* 7543 8661: contig of 1119 bp in length  
\* 8662 8761: gap of unknown length  
\* 8762 10030: contig of 1269 bp in length  
\* 10031 10130: gap of unknown length  
\* 10131 11491: contig of 1361 bp in length  
\* 11492 11591: gap of unknown length  
\* 11592 12867: contig of 1376 bp in length  
\* 12868 13067: gap of unknown length  
\* 13068 14173: contig of 1106 bp in length  
\* 14174 14273: gap of unknown length  
\* 14274 15840: contig of 1567 bp in length  
\* 15841 15940: gap of unknown length  
\* 15941 17242: contig of 1302 bp in length  
\* 17243 17342: gap of unknown length  
\* 17343 18449: contig of 1107 bp in length  
\* 18450 18549: gap of unknown length  
\* 18550 19683: contig of 1134 bp in length  
\* 19684 19783: gap of unknown length  
\* 19784 21247: contig of 1464 bp in length  
\* 21248 21347: gap of unknown length  
\* 21348 22952: contig of 1605 bp in length  
\* 22953 23052: gap of unknown length  
\* 23053 24452: contig of 1400 bp in length  
\* 24453 24552: gap of unknown length  
\* 24553 25825: contig of 1273 bp in length  
\* 25826 25925: gap of unknown length  
\* 25926 27109: contig of 1184 bp in length  
\* 27110 28292: gap of unknown length  
\* 28293 28392: gap of unknown length  
\* 28393 29744: contig of 1352 bp in length  
\* 29745 29844: gap of unknown length  
\* 29845 31042: contig of 1198 bp in length  
\* 31043 31142: gap of unknown length  
\* 31143 32629: contig of 1467 bp in length  
\* 32630 32729: gap of unknown length  
\* 32730 33698: contig of 1169 bp in length  
\* 33699 33989: gap of unknown length  
\* 33990 35704: contig of 1706 bp in length  
\* 35705 35804: gap of unknown length  
\* 35805 37471: contig of 1667 bp in length  
\* 37472 37571: gap of unknown length  
\* 37572 39285: contig of 1714 bp in length  
\* 39286 39385: gap of unknown length  
\* 39386 40577: contig of 1192 bp in length  
\* 40578 40677: gap of unknown length  
\* 40679 41930: contig of 1253 bp in length  
\* 41931 42030: gap of unknown length  
\* 42031 43743: contig of 1713 bp in length  
\* 43744 43843: gap of unknown length  
\* 43844 46054: contig of 2211 bp in length  
\* 46055 46154: gap of unknown length  
\* 46155 49379: contig of 3225 bp in length  
\* 49380 49479: gap of unknown length  
\* 49480 53371: contig of 3892 bp in length  
\* 53372 53471: gap of unknown length  
\* 53472 55756: contig of 2285 bp in length  
\* 55757 55856: gap of unknown length  
\* 55857 58019: contig of 2163 bp in length  
\* 58020 58119: gap of unknown length  
\* 58120 60327: contig of 2208 bp in length  
\* 60328 60427: gap of unknown length  
\* 60429 64223: contig of 3796 bp in length  
\* 64224 64323: gap of unknown length  
\* 64324 67218: contig of 2895 bp in length  
\* 67219 67318: gap of unknown length  
\* 67319 71797: contig of 4479 bp in length  
\* 71798 71897: gap of unknown length  
\* 71898 76856: contig of 4959 bp in length  
\* 76857 76956: gap of unknown length

\* 76957 82060: contig of 5104 bp in length  
\* 82061 82160: gap of unknown length  
\* 82161 87557: contig of 5397 bp in length  
\* 87558 87657: gap of unknown length  
\* 87658 94226: contig of 6569 bp in length  
\* 94227 94326: gap of unknown length  
\* 94327 101878: contig of 7552 bp in length  
\* 101879 101978: gap of unknown length  
\* 101979 110629: contig of 8651 bp in length  
\* 110630 110729: gap of unknown length  
\* 110730 119083: contig of 8354 bp in length  
\* 119084 119183: gap of unknown length  
\* 119184 128919: contig of 9736 bp in length  
\* 128920 129019: gap of unknown length  
\* 129020 142791: contig of 13772 bp in length  
\* 142792 142891: gap of unknown length  
\* 142892 156079: contig of 13188 bp in length  
\* 156080 156179: gap of unknown length  
\* 156180 183092: contig of 26913 bp in length  
\* 183093 183192: gap of unknown length  
\* 183193 216441: contig of 33249 bp in length.  
Location/Qualifiers  
1. 216441  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clone="CTD-2257C19"  
/clone\_1fb="Caltech human BAC library D"  
1042. 1141  
/estimated\_length=unknown  
gap 2371. 2470  
/estimated\_length=unknown  
gap 3553. 3652  
/estimated\_length=unknown  
gap 4684. 4783  
/estimated\_length=unknown  
gap 6179. 6278  
/estimated\_length=unknown  
gap 7443. 7542  
/estimated\_length=unknown  
gap 8662. 8761  
/estimated\_length=unknown  
gap 10031. 10130  
/estimated\_length=unknown  
gap 11492. 11591  
/estimated\_length=unknown  
gap 12968. 13067  
/estimated\_length=unknown  
gap 14174. 14273  
/estimated\_length=unknown  
gap 15841. 15940  
/estimated\_length=unknown  
gap 17243. 17342  
/estimated\_length=unknown  
gap 18450. 18549  
/estimated\_length=unknown  
gap 19684. 19783  
/estimated\_length=unknown  
gap 21248. 21347  
/estimated\_length=unknown  
gap 22953. 23052  
/estimated\_length=unknown  
gap 24453. 24552  
/estimated\_length=unknown  
gap 25826. 25925  
/estimated\_length=unknown  
gap 27110. 27209  
/estimated\_length=unknown  
gap 28293. 28392  
/estimated\_length=unknown  
gap 29745. 29844

Query Match 12.6%; Score 392; DB 14; Length 216441;  
Best Local Similarity 99.6%; Pred. No. 3,5e-202;  
Matches 752; Conservative 0; Mismatches 0; Indels 3; Gaps 3;  
QY 2036 TGCATCAAAATGTCAGCAGCTTTTATTAATCTTAATCTTTACAGGGCTTAATTTAGAGAG 2095  
DB TGCAATCAAAATGTCAGCAGCTTTTATTAATCTTAATCTTTACAGGGCTTAATTTAGAGAG 42347  
QY 2096 TGTCTGAGAGAGTTCAATCAAAAGGCTTTCTTAAGACCGCTCAACGCCCTTCTAGC 2155  
DB TGTCTGAGAGAGTTCAATCAAAAGGCTTTCTTAAGACCGCTTCAACCGCTTCTAGC 42348  
QY 2156 AGAGTTATTCATTCGTCGCCAAGAGAGCTTGAAGAGATTGAAGTATGATGACCTCCAC 2215  
DB AGAGTTATTCATTCGTCGCCAAGAGAGCTTGAAGAGATTGAAGTATGATGACCTCCAC 42408  
QY 2216 TCCGCTCAGAGGAGCTGACCTTATTTAGAAAACAAAGAGGCTGAGTGAACCTACTCA 2275  
DB TCCGCTCAGAGGAGCTGACCTTATTTAGAAAACAAAGAGGCTGAGTGAACCTACTCA 42467  
QY 2276 CGAATTCGATTCAGATGCGCACTTGCTGCGGAAAAGGCTCTCCACAGCCCGGA 2335  
DB CGAATTCGATTCAGATGCGCACTTGCTGCGGAAAAGGCTCTCCACAGCCCGGA 42467  
QY 2336 CTTTCTCCTCTCTGAGGCTTATTTTGTGAGACTAGACAGAGTGTGAACTTCT 2455  
DB CTTTCTCCTCTCTGAGGCTTATTTTGTGAGACTAGACAGAGTGTGAACTTCT 42647  
QY 2456 TTGCAGAGAGGCTGAGAACTCTTTAGAGCACTTAATCTTAATTTTCCCTGGAATGTG 2515  
DB TTGCAGAGAGGCTGAGAACTCTTTAGAGCACTTAATCTTAATTTTCCCTGGAATGTG 42706  
QY 2516 CGTGTGCGCAGTAAAGAGGCGCTTGGAGAGCTCCCTGACCCCGCGCTGCCCGCC 2575  
DB CGTGTGCGCAGTAAAGAGGCGCTTGGAGAGCTCCCTGACCCCGCGCTGCCCGCC 42766  
QY 2576 TCCGCGGTATGTGCACTTAATCTGCGCCACAGAGGTTTGAAGCAATCAAGCTGAGACTG 2635  
DB TCCGCGGTATGTGCACTTAATCTGCGCCACAGAGGTTTGAAGCAATCAAGCTGAGACTG 42825  
QY 2636 GGTGAATGTAAACGCTTTAATCTTGAAGTTTAAGAACTTTAAAGCTTAATCTCTC 2695  
DB GGTGAATGTAAACGCTTTAATCTTGAAGTTTAAGAACTTTAAAGCTTAATCTCTC 42885  
QY 2696 TGAAGAAAATGACGTAAACAGAGCTGTACTATGAAGCTGTTATTTAATAAGAAC 2755  
DB TGAAGAAAATGACGTAAACAGAGCTGTACTATGAAGCTGTTATTTAATAAGAAC 42945  
QY 2756 GCTGGGCAATGAATCAATCACTGCGCAATGAGTCAA 2790  
DB GCTGGGCAATGAATCAATCACTGCGCAATGAGTCAA 43005  
RESULT 12  
G37408/c 449 bp DNA linear STS 31-MAR-1998  
LOCUS SHGC-57769 Human Homo sapiens STS genomic, sequence tagged site.  
DEFINITION G37408  
ACCESSION G37408  
VERSION G37408.1 GI:2997059  
KEYWORDS STS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 449)  
AUTHORS Myers, R.M.  
TITLE Human STS (1997)  
JOURNAL Unpublished (1997)

## COMMENT

Contact: Richard M. Myers  
Stanford Human Genome Center (SHGC)  
Stanford University School of Medicine  
Department of Genetics, W-344, Stanford, CA 94305, USA  
Tel: 4157259687  
Fax: 4157259689  
Email: myers@hgc.stanford.edu  
Primer A: CATGCCGAGCTCTTAT  
Primer B: ACTGCCCCACAGAGCTTTG  
STS size: 173  
PCR Profile:  
Initial incubation: 95 degrees C for 10 minutes

Denaturation: 94 degrees C for 30 seconds  
Annealing: 60 degrees C for 30 seconds  
Polymerization: 72 degrees C for 23 seconds  
PCR Cycles: 30  
Thermal Cycler: Perkin Elmer 9700  
Protocol:  
Template: 25 ng  
Primer: each 1 uM  
dNTPs: each 200 uM  
AmpliTaq Gold Polymerase: 0.07 units/ul  
Total Vol: 5 ul

Buffer:  
MgCl2: 2.5 mM  
KCl: 50 mM  
Tris-HCl: 10 mM  
pH: 8.3

Prepared with primer pairs derived from W81222 -- Unigene.

## FEATURES

source  
1..449  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/map="19"  
/clone\_lib="Human"  
STS  
primer\_bind  
primer\_bind  
complement(163..182)

Query Match 4.6%; Score 144; DB 10; Length 449;  
Best Local Similarity 100.0%; Pred. No. 2.1e-66;  
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2631 GACTGGCTTGAATGTAAGCTTTAACTTGAAGCTTTAAAGCTTAATA 2690  
DB 145 GACTGGCTTGAATGTAAGCTTTAACTTGAAGCTTTAAAGCTTAATA 86  
QY 2691 TCCTCGAAGAAAATGACGTACACAGCGTGTACTATGAAGCTTTAATTATA 2750  
DB 85 TCCTCGAAGAAAATGACGTACACAGCGTGTACTATGAAGCTTTAATTATA 26  
QY 2751 AGAAGCTGGGCCATGAACCTCAT 2774  
DB 25 AGAAGCTGGGCCATGAACCTCAT 2  
RESULT 13  
CQ553695  
LOCUS CQ553695 60 bp DNA linear PAT 30-JAN-2004  
DEFINITION Sequence 23330 from Patent WO0210449.  
ACCESSION CQ553695  
VERSION CQ553695.1 GI:41520122  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

## REFERENCE

1 Shoshan,A., Wasserman,A., Mintz,E., Mintz,L. and Paigler,S.  
TITLE Oligonucleotide library for detecting rna transcripts and splice  
JOURNAL variants that populate a transcritpome  
Patent: WO 0210449-A 23330 07-FEB-2002;  
CompuGen Inc. (US)

## FEATURES

source  
1..60  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"

## ORIGIN

Query Match 1.9%; Score 60; DB 6; Length 60;  
Best Local Similarity 100.0%; Pred. No. 2e-20;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2464 GGGCTGGGATCTCTTTAGACACTTAATCTATTTATCCCTGGATGCGTCTGG 2523  
DB 1 GGGCTGGGATCTCTTTAGACACTTAATCTATTTATCCCTGGATGCGTCTGG 60

RESULT 14  
CR762434 24495 bp DNA linear PRI 16-FEB-2005  
LOCUS CR762434  
DEFINITION Human DNA sequence from clone DADB-91M20 on chromosome 6, complete  
sequence.  
ACCESSION CR762434  
VERSION CR762434.7 GI:59891252  
KEYWORDS HTG.

## SOURCE

Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 24495)  
Almeida,J.

## REFERENCE

AUTHORS Submitted (16-FEB-2005) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquerry@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk  
On Feb 16, 2005 this sequence version replaced gi:56369811.  
JOURNAL  
TITLE  
COMMENT

Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk)

During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest, except on the rare  
occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information  
on the WORMPEP database can be found at

[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence  
was generated from part of bacterial clone configs of human  
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping  
Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr6>  
DADB-91M20 is from a DNA-arts DBB human bac library VECTOR:



FEATURES pBelOBAC11.  
source 1.2435 Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone\_lib="DNA-arts-BAC.1-DBB.1"

ORIGIN  
Query Match 1.7%; Score 53; DB 8; Length 24495;  
Best Local Similarity 100.0%; Pred. No. 1.6e-16;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2888 TGAGGAGGTGATCACTGAGGCCAGAGTTGAGAGCCAGCTGGCCAACT 2940  
DB 8493 TGAGGAGGTGATCACTGAGGCCAGAGTTGAGAGCCAGCTGGCCAACT 8545

RESULT 15  
DB4401/c 42398 bp DNA linear PRI 20-NOV-1999  
LOCUS Homo sapiens genomic DNA, 43 kb segment from chromosome6, complete  
DEFINITION sequence.  
ACCESSION DB4401 GI:2401263  
VERSION DB4401.1  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases)  
AUTHORS Kikuti,Y.Y., Tamiya,G., Ando,A., Chen,L., Kimura,M., Ferreira,E.,  
Teuji,K., Trowdale,J., and Inoko,H.  
TITLE Physical mapping 220 kb centromeric of the human MHC and DNA  
sequence analysis of the 43-kb segment including the RING1, HKE6,  
and HKE4 genes  
JOURNAL Genomics 42 (3), 422-435 (1997)  
PUBMED 9205114  
REFERENCE 2 (bases 1 to 42398)  
AUTHORS Kikuti,Y., Inoko,H., Ando,A., Kimura,M., Watanabe,K. and Shina,T.  
TITLE Physical map of 200 kb at the centromeric side of the human MHC  
region and sequence of 43 kb  
JOURNAL Unpublished  
REFERENCE 3 (bases 1 to 42398)  
AUTHORS Inoko,H.  
TITLE Direct Submision  
JOURNAL Submitted (12-APR-1996) Hidetoshi Inoko, Tokai University School of  
Medicine, Molecular Life Science; Bohseidai, Isehara, Kanagawa  
259-11, Japan (E-mail:hino@mls.lsc.u-tokai.ac.jp,  
Tel:0463-93-1121, Fax:0463-94-8884)  
Location/Qualifiers  
FEATURES source 1.42398  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone\_lib="519"  
/cell\_line="CGMT"

ORIGIN  
Query Match 1.7%; Score 53; DB 8; Length 42398;  
Best Local Similarity 100.0%; Pred. No. 1.6e-16;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2888 TGAGGAGGTGATCACTGAGGCCAGAGTTGAGAGCCAGCTGGCCAACT 2940  
DB 5305 TGAGGAGGTGATCACTGAGGCCAGAGTTGAGAGCCAGCTGGCCAACT 5253

RESULT 16

CR547129  
LOCUS CR547129 49878 bp DNA linear PRI 18-JAN-2005  
DEFINITION Human DNA sequence from clone DANA-258A19 on chromosome 6, complete  
sequence.  
ACCESSION CR547129  
VERSION CR547129.3 GI:57898954  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 49878)  
AUTHORS Wood,J.  
TITLE Direct Submision  
JOURNAL Submitted (18-JAN-2005) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk  
On Jan 18, 2005 this sequence version replaced gi:5029847.  
COMMENT ----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
-----  
During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest, except on the rare  
occasion of the clone being a YAC.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information  
on the WORMPEP database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep This sequence  
was generated from part of bacterial clone contigs constructed by  
the MHC HaploType Consortium and collaborators. Further information  
can be found at  
http://www.sanger.ac.uk/HGP/Chr6/MHC  
DANA-258A19 is from the DNA-Arts human BAC library MANN.1 VECTOR:  
pBelOBAC11.  
Location/Qualifiers  
FEATURES source 1.49878  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone\_lib="DANA-258A19"  
/clone\_lib="DNA-Arts BAC library MANN.1"

ORIGIN  
Query Match 1.7%; Score 53; DB 8; Length 49878;  
Best Local Similarity 100.0%; Pred. No. 1.6e-16;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2888 TGAGGAGGTGATCACTGAGGCCAGAGTTGAGAGCCAGCTGGCCAACT 2940  
DB 30123 TGAGGAGGTGATCACTGAGGCCAGAGTTGAGAGCCAGCTGGCCAACT 30175

RESULT 17  
AC061985 62449 bp DNA linear HTG 21-APR-2000  
LOCUS AC061985  
DEFINITION Homo sapiens chromosome 17 clone -2009115 map 17, LOW-PASS SEQUENCE

SAMPLING.  
AC061985  
AC061985.1 GI:7630724  
HTG; HTGS\_PHASE0.  
KEYWORDS  
Homo sapiens (human)  
SOURCE  
Homo sapiens  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
1 (bases 1 to 62449)  
Birren,B., Linton,L., Nusbaum,C. and Lander,B.  
Homo sapiens chromosome 17, clone -2009115  
Unpublished  
2 (bases 1 to 62449)  
Birren,B., Linton,L., Nusbaum,C., Lander,B., Abraham,H., Allen,N.,  
Anderson,S., Baldwin,J., Barna,N., Baeten,V., Beda,F.,  
Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G.,  
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,  
Collymore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,  
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,  
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
Grand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,  
Howard,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
Klein,J., Laroque,K., Lamazares,R., Landers,T., Lehoczy,J.,  
Levine,R., Lieu,C., Liu,G., Locke,K., MacDonald,P., Marquis,N.,  
McCarthy,M., McEwan,P., McGurk,A., McKernan,K., McPheters,R.,  
Meldrum,J., Menes,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,  
Murphy,T., Naylor,T., Norman,C.H., O'Connor,T., O'Donnell,P.,  
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,  
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
Roy,A., Santos,R., Schoenert,S., Severy,P., Spencer,B.,  
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Testaye,S., Theodore,J., Tittel,A., Travers,M., Ttigilio,J.,  
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
Young,G., Zainoun,J., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (21-APR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
All repeats were identified using RepeatMasker:  
Smit,A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence.submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L9672  
Center clone name: 2009\_I\_15  
-----  
\* NOTE: This record contains 72 individual  
\* sequencing reads that have not been assembled into  
\* contigs. Runs of N are used to separate the reads  
\* and the order in which they appear is completely  
\* arbitrary. Low-pass sequence sampling is useful for  
\* identifying clones that may be gene-rich and allows  
\* overlap relationships among clones to be deduced.  
\* However, it should not be assumed that this clone  
\* will be sequenced to completion. In the event that  
\* the record is updated, the accession number will  
\* be preserved.  
\* 1 753: contig of 753 bp in length  
\* 754 853: gap of 100 bp  
\* 854 1619: contig of 766 bp in length  
\* 1620 1719: gap of 100 bp  
\* 1720 2489: contig of 770 bp in length  
\* 2490 2589: gap of 100 bp  
\* 2590 3157: contig of 776 bp in length  
\* 3158 3467: gap of 100 bp  
\* 3468 4239: contig of 772 bp in length  
\* 4240 4339: gap of 100 bp  
\* 4340 5105: contig of 766 bp in length  
\* 5106 5205: gap of 100 bp  
\* 5206 5967: contig of 762 bp in length  
\* 5968 6067: gap of 100 bp  
\* 6068 6846: contig of 779 bp in length  
\* 6847 6946: gap of 100 bp  
\* 6947 7717: contig of 771 bp in length  
\* 7718 7817: gap of 100 bp  
\* 7818 8586: contig of 769 bp in length  
\* 8587 8686: gap of 100 bp  
\* 8687 9488: contig of 802 bp in length  
\* 9489 9588: gap of 100 bp  
\* 9589 10353: contig of 765 bp in length  
\* 10354 10453: gap of 100 bp  
\* 10454 11229: contig of 776 bp in length  
\* 11230 11329: gap of 100 bp  
\* 11330 12086: contig of 757 bp in length  
\* 12087 12186: gap of 100 bp  
\* 12187 12958: contig of 772 bp in length  
\* 12959 13058: gap of 100 bp  
\* 13059 13841: contig of 783 bp in length  
\* 13842 13941: gap of 100 bp  
\* 13942 14700: contig of 759 bp in length  
\* 14701 14800: gap of 100 bp  
\* 14801 15557: contig of 757 bp in length  
\* 15558 15657: gap of 100 bp  
\* 15658 16431: contig of 774 bp in length  
\* 16432 16531: gap of 100 bp  
\* 16532 17306: contig of 775 bp in length  
\* 17307 17406: gap of 100 bp  
\* 17407 18178: contig of 772 bp in length  
\* 18179 18788: gap of 100 bp  
\* 18789 19032: contig of 754 bp in length  
\* 19033 19132: gap of 100 bp  
\* 19133 19883: contig of 751 bp in length  
\* 19884 19983: gap of 100 bp  
\* 19984 20761: contig of 778 bp in length  
\* 20762 20861: gap of 100 bp  
\* 20862 21636: contig of 775 bp in length  
\* 21637 21737: gap of 100 bp  
\* 21737 22520: contig of 784 bp in length  
\* 22521 22620: gap of 100 bp  
\* 22621 23398: contig of 778 bp in length  
\* 23399 23498: gap of 100 bp  
\* 23499 24277: contig of 779 bp in length  
\* 24278 24377: gap of 100 bp  
\* 24378 25163: contig of 786 bp in length  
\* 25164 25263: gap of 100 bp  
\* 25264 26026: contig of 763 bp in length  
\* 26027 26126: gap of 100 bp  
\* 26127 26892: contig of 766 bp in length  
\* 26893 26992: gap of 100 bp  
\* 26993 27753: contig of 761 bp in length  
\* 27754 27853: gap of 100 bp  
\* 27854 28621: contig of 768 bp in length  
\* 28622 28721: gap of 100 bp  
\* 28722 29499: contig of 778 bp in length  
\* 29500 29599: gap of 100 bp  
\* 29600 30370: contig of 771 bp in length  
\* 30371 30470: gap of 100 bp  
\* 30471 31247: contig of 777 bp in length  
\* 31248 31347: gap of 100 bp  
\* 31348 32118: contig of 771 bp in length  
\* 32119 32218: gap of 100 bp  
\* 32219 32946: contig of 728 bp in length  
\* 32947 33046: gap of 100 bp  
\* 33047 33824: contig of 778 bp in length  
\* 33825 33924: gap of 100 bp  
\* 33925 34678: contig of 754 bp in length  
\* 34679 34778: gap of 100 bp  
\* 34779 35551: contig of 773 bp in length  
\* 35552 35651: gap of 100 bp  
\* 35652 36422: contig of 771 bp in length  
\* 36423 36522: gap of 100 bp  
\* 36523 37295: contig of 773 bp in length

```
* 37296 37296: gap of 100 bp
* 37396 38150: contig of 765 bp in length
* 38161 38260: gap of 100 bp
* 38261 39025: contig of 765 bp in length
* 39026 39125: gap of 100 bp
* 39126 39879: contig of 754 bp in length
* 39880 39979: gap of 100 bp
* 39980 40746: contig of 767 bp in length
* 40747 40846: gap of 100 bp
* 40847 41620: contig of 774 bp in length
* 41621 41720: gap of 100 bp
* 41721 42483: contig of 763 bp in length
* 42484 42583: gap of 100 bp
* 42584 43359: contig of 776 bp in length
* 43360 43459: gap of 100 bp
* 43460 44236: contig of 777 bp in length
* 44237 44336: gap of 100 bp
* 44337 45111: contig of 775 bp in length
* 45112 45211: gap of 100 bp
* 45212 45981: contig of 770 bp in length
* 45982 46081: gap of 100 bp
* 46082 46844: contig of 763 bp in length
* 46845 46944: gap of 100 bp
* 46945 47709: contig of 765 bp in length
* 47710 47809: gap of 100 bp
* 47810 48550: contig of 741 bp in length
* 48551 48650: gap of 100 bp
* 48651 49420: contig of 770 bp in length
* 49421 49520: gap of 100 bp
* 49521 50284: contig of 764 bp in length
* 50285 50384: gap of 100 bp
* 50385 51169: contig of 785 bp in length
* 51170 51269: gap of 100 bp
* 51270 52054: contig of 785 bp in length
* 52055 52154: gap of 100 bp
* 52155 52930: contig of 776 bp in length
* 52931 53030: gap of 100 bp
* 53031 53803: contig of 773 bp in length
* 53804 53903: gap of 100 bp
* 53904 54670: contig of 767 bp in length
* 54671 54770: gap of 100 bp
* 54771 55546: contig of 776 bp in length
* 55547 55646: gap of 100 bp
* 55647 56408: contig of 762 bp in length
* 56409 57273: gap of 100 bp
* 57274 57373: contig of 765 bp in length
* 57374 58141: gap of 100 bp
* 58142 58241: gap of 100 bp
* 58242 58983: contig of 742 bp in length
* 58984 59084: gap of 100 bp
* 59084 59851: contig of 768 bp in length
```

Query Match 1.7%; Score 53; DB 14; Length 62449;

Best Local Similarity 100.0%; Pred. No. 1.6e-16;

Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2888 TGAGGAGGTGATCCTGAGGCCGAGAGTTGAGACCAAGCTGGCCCAACT 2940  
|||||  
DB 1892 TGAGGAGGTGATCCTGAGGCCGAGAGTTGAGACCAAGCTGGCCCAACT 1944

RESULT 19  
AC026771/c AC026771 69808 bp DNA linear PRI 13-MAR-2003  
LOCUS Homo sapiens chromosome 16 clone CTB-134F13, complete sequence.  
AC026771 AC026771  
VERSION AC026771.6 GI:28933549  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
UNPUBLISHED

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
UNPUBLISHED

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
UNPUBLISHED

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
UNPUBLISHED

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
UNPUBLISHED

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
UNPUBLISHED

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
UNPUBLISHED

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
UNPUBLISHED

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
UNPUBLISHED

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
UNPUBLISHED

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
UNPUBLISHED

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
UNPUBLISHED

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
UNPUBLISHED

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
UNPUBLISHED

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
UNPUBLISHED

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
UNPUBLISHED

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catherhini;  
Homnidae; Homo.  
1 (bases 1 to 70621)  
Almeida, J.  
Direct Submission  
Submitted (17-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,

----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [vega@sanger.ac.uk](mailto:vega@sanger.ac.uk)

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

```

/note="match: ESTs: AI186776.1 AI359146.1 AI992249.1
AM134912.1 BQ218704.1"
join(AL662824.9:186288.186339,AL662824.9:186521.186738
CDS

```

```

Gene      join14451..4495,4416..4451,5277..5437,5566..5781,
          6803..7198,7280..7653,7949..8333)
          /gene="RING1"
locus_tag="XXbac-BCX269C15.3..002"
join1411..4295,4416..4451,5277..5437,5566..5781,
rRNA

```

**CDS**

gene

mRNA

**CDS**

**mRNA**

**gene**

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

**COMMENT**

## FEATURES

Source

misc\_feature

**gene**

**mRNA**

```

/protein_id="CA117651.1"
/db_xref="GI:55961768"
/db_xref="InterPro:IPR001841"
/location="MDGTEIAVSPRLSHSELMCPICLMDLKNMTTKECLHFCSDCI
VIALSGNKECTCRKLVSKRSLRDPDPFDLILKIPSRREYAHODRVLIRSL
HNOALSSIEGELKLMQAMHRAQVRPPIPGSDOTTSAGSEDEGDEJEDVSD
SAPDAPGPAPKRPBGAGGAGSVGTGGGTGCGVAGGSEDSGRGTIGGGTIGGP
SPGAPSPPEPGGRTIYVPRPILVKGECOTPVYTTGNATYDMLSKYATAPAL
ERRQOQAGEPGPGGASDTGPGGCGGAGGAGGDPPEPALPSLEGSEKQYTI
YIAPGGAGTTTNGSLITLVLNKKFKVSRPLKCYAPTKDKR"
gene
11366. 11973
/locus_tag="XXbac-BCX269C15.2-001"
/pseudo
11366. 11973
/locus_tag="XXbac-BCX269C15.2-001"
/locus="match: proteinins: P28698 Q96H95 Q96I71 Q9NRY0"
/locus="match: proteinins: P28698 Q96H95 Q96I71 Q9NRY0"
/pseudo
/codon_start=1
/product="zinc finger protein 314 pseudogene"
join(33490..34019,34322..34431,35078..35432)
/locus_tag="XXbac-BCX269C15.1-003"
/locus="match: proteinins: O43719 Q99730 Q9CT41 Q9DAU3"
/pseudo
/codon_start=1
/product="HIV TAT specific factor 1 pseudogene"
join(45181..45302,45437..45536,46263..46358,47160..47353,
50345..50531)
/locus_tag="XXbac-BCX269C15.5-001"
join(45181..45302,45437..45536,46263..46358,47160..47353,
50345..50531)
/locus_tag="XXbac-BCX269C15.5-002"
join(45181..45302,45437..45536,46263..46358,47160..47353,
50345..50531)
gene
mrna
Query Match 1.7%; Score 53; DB 8; Length 70621;
Best Local Similarity 100.0%; Pred. No. 1.6e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2888 TGAGGAGGTGATCACTGAGCGCAGAGTTTCAGACACGAGCGTGGCAACAT 2940
DB 34070 TGAGGAGGTGATCACTGAGCGCAGAGTTTCAGACACGAGCGTGGCAACAT 34122
RESULT 20 90476 bp DNA linear PRI 01-APR-2002
AC015910
LOCUS Homo sapiens chromosome 17, clone RP11-690G19, complete sequence.
AC015910
DEFINITION AC015910.10 GI:19852158
VERSION AC015910.10
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 90476)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N., Anderson,M.,
Baldwin,J., Barina,N., Beckerly,R., Boguslavsky,L., Bouhgalter,B.,
Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A.,
Cooke,P., Dearellano,K., Dewar,K., Donlin,M., Donejan,L., Doyle,M.,
Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D.,
Galagan,J., Gardyna,S., Grant,G., Hages,B., Heaford,A., Horton,L.,
Howard,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,U.,
Lehoczky,J., Liu,C., Locke,K., Macdonald,P., Margulis,N.,
McEwan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J.,
Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,

```

Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P.,  
 Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
 Tesfaye,S., Tittel,A., Vassiliev,H., Vo,A., Wheeler,J., Wu,X.,  
 Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.

Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 90476)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N.,  
 Anderson,S., Barina,N., Baetien,V., Bloom,T., Boguslavsky,L.,  
 Bouhgalter,B., Brown,A., Camarata,C., Campopiano,A., Chang,J.,  
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,  
 Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
 Fero,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,  
 Ginde,S., Gord,S., Goyette,M., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., Larcocque,K., Lamazares,R.,  
 Lander,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,  
 Maclean,C., Macdonald,P., Major,J., Margulis,N., Matthews,C.,  
 McCarty,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,  
 Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,  
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schubovic,N.,  
 Seaman,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N.,  
 Strause,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
 Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,D., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

#### Direct Submission

Submitted (22-MAR-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 4 (bases 1 to 90476)

Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N.,  
 Anderson,S., Barina,N., Baetien,V., Bloom,T., Boguslavsky,L.,  
 Bouhgalter,B., Brown,A., Camarata,C., Campopiano,A., Chang,J.,  
 Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A.,  
 Cook,A., Cooke,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,  
 Fero,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,  
 Ginde,S., Gord,S., Goyette,M., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
 Kamat,A., Karatas,A., Kells,C., Larcocque,K., Lamazares,R.,  
 Lander,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,  
 Maclean,C., Macdonald,P., Major,J., Margulis,N., Matthews,C.,  
 McCarty,M., McEwan,P., McKernan,K., Meldrum,J., Meneus,L.,  
 Mihova,T., Mienga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,  
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
 Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
 Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
 Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schubovic,N.,  
 Seaman,S., Severy,P., Spencer,B., Strange-Thomann,N., Stojanovic,N.,  
 Strause,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,  
 Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H.,  
 Viel,R., Vo,A., Wilson,D., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
 Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

Submitted (01-APR-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Apr 1, 2002 this sequence version replaced gi:19852158.

All repeats were identified using RepeatMasker:  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center  
 Center: Whitehead Institute/MIT Center for Genome Research

Center code: WMRB  
 Web site: http://www-seq.wi.mit.edu

Contact: sequence submissions@genome.wi.mit.edu  
 Project Information

Center project name: I632  
 Center clone name: 690\_G\_19

Only the first 90.5 kilobases of this clone are being submitted.  
The remainder of the clone is overlapped by either accession number  
AC005288 [WICR project L301] or accession  
number AC004408 [WICR project L309].

## FEATURES

source  
1..90476  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/map="17"  
/clone="RP11-690G19"  
/clone\_lib="RP11-690G19 Human Male BAC"  
111..405  
repeat\_region  
/rpt\_family="AluB"  
complement(572..750)  
/rpt\_family="AluSg/x"  
complement(752..920)  
/rpt\_family="AluSx"  
complement(921..1140)  
/rpt\_family="AluSg/x"  
complement(1141..1267)  
/rpt\_family="AluSx"  
complement(1278..1559)  
/rpt\_family="AluB"  
complement(1660..1947)  
/rpt\_family="AluB"  
complement(2011..2302)  
/rpt\_family="AluSx"  
2330..2625  
/rpt\_family="AluSx"  
2378..2386  
/note="extensive edit here"  
2749..3024  
/rpt\_family="AluSx"  
3049..3126  
/rpt\_family="CATADA)n"  
complement(3393..3699)  
/rpt\_family="AluSg"  
complement(4229..4354)  
/rpt\_family="FLAM\_A"  
4471..4521  
/rpt\_family="Alu"  
complement(4522..4785)  
/rpt\_family="AluSg"  
4793..4919  
/rpt\_family="TTCC)n"  
complement(4931..5209)  
/rpt\_family="AluSx"  
5465..5772  
/rpt\_family="AluSx"  
complement(6691..6796)  
/rpt\_family="FLAM\_C"  
complement(6833..6928)  
/rpt\_family="AluB"  
complement(6963..6999)  
/rpt\_family="MIR"  
7000..7287  
/rpt\_family="AluB"  
complement(7288..7338)  
/rpt\_family="MIR"  
complement(7339..7628)  
/rpt\_family="AluSg"  
complement(7629..7752)  
/rpt\_family="MIR"  
7859..8138  
/rpt\_family="AluSx"  
complement(8415..8748)  
/rpt\_family="AluSx"  
complement(8772..9067)  
/rpt\_family="AluSg"  
complement(9082..9207)  
/rpt\_family="FLAM\_C"

repeat\_region complement(9550..9858)  
/rpt\_family="AluSx"  
repeat\_region complement(10613..10908)  
/rpt\_family="AluSx"  
repeat\_region 11112..11147  
/rpt\_family="AT\_rich"  
repeat\_region 11261..11376  
/rpt\_family="AluB"  
repeat\_region 11377..11681  
/rpt\_family="AluSg"  
repeat\_region 11682..11996  
/rpt\_family="AluB"  
repeat\_region 11997..12164  
/rpt\_family="AluB"  
complement(12789..13093)  
/rpt\_family="AluSg"  
complement(13819..13888)  
/rpt\_family="AluB/FRAM"  
complement(13889..14203)  
/rpt\_family="AluSx"  
complement(14204..14303)  
/rpt\_family="AluB/FRAM"  
repeat\_region 14528..14656  
/rpt\_family="AluSx"  
repeat\_region 14657..14957  
/rpt\_family="AluSx"  
repeat\_region 15137..15372  
/rpt\_family="AluB"  
repeat\_region 15389..15657

Query Match 1.7%; Score 53; DB 8; Length 90476;  
Best Local Similarity 100.0%; Pred.No. 1.6e-16;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2888 TGAGGACAGTGTGATCACTGAGGCCAGAGTTGAGACCGCTGCGCAACAT 2940  
Db 14575 TGAGGACAGTGTGATCACTGAGGCCAGAGTTGAGACCGCTGCGCAACAT 14627  
|||||  
|||||

RESULT 21  
AP000208 100000 bp DNA linear PRI 08-JAN-2000  
LOCUS Homo sapiens genomic DNA, chromosome 21q22.1, D21S226-AML region,  
DEFINITION Clone f43D11-11988, segment 6/12, complete sequence.  
ACCESSION AP000208  
VERSION AP000208.1 GI:4827146  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 100000)  
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,  
Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.  
Homo sapiens 1,109,292bp genomic DNA of 21q22.1 (REGION:  
D21S226-AML CLONE RANGE: f43D11-11988)  
2 (bases 1 to 100000)  
Published Only in Database (1999)  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
DIRECT SUBMISSION  
Fujiyama,A., Yada,T., Totoki,Y. and Sakaki,Y.  
Submitted (10-MAY-1999) Maahira Hattori, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),  
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,  
Japan (E-mail:hattori@gsc.riken.go.jp/  
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,  
Fax:81-42-778-9924)  
E. coli transposon insertion: The present data does not contain E.  
coli transposon sequences which integrated in the  
original/previous sequences. We determined the boundary between  
the insertion and genomic sequences experimentally, removed the  
insertion sequences, reconstituted the present data. The sequencing

## COMMENT



project is supported by Japan Science Technology Corporation (JST)  
and The Institute of Physical and Chemical Research (RIKEN).  
Location/Qualifiers

## FEATURES

## source

1. 100000  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.1"

## ORIGIN

Query Match 1.7%; Score 53; DB 8; Length 100000;  
Best Local Similarity 100.0%; Pred. No. 1.6e-16;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2888 TGAGCAGGTGATGATCACTGAGCCAGAGATTGAGACCAAGCTGCGCAACAT 2940  
3384 TGAGCAGGTGATGATCACTGAGCCAGAGATTGAGACCAAGCTGCGCAACAT 33436

RESULT 22  
LOCUS AL844527 109523 bp DNA linear PRI 18-MAY-2005  
DEFINITION Human DNA sequence from clone DAOB-314F24 on chromosome 6. Contains  
the 5' end of the COL11A2 gene for collagen type XI alpha 2, the  
RXR gene for retinoid X receptor, beta, the SLC39A7 gene for  
soluble carrier family 39 (zinc transporter), member 7, the HSD17B8  
gene for hydroxysteroid (17-beta) dehydrogenase 8, the RING1 gene  
for ring finger protein 1, the ZNF314P pseudogene for zinc finger  
protein 314, the HTRAF1P pseudogene for HIV TAT specific factor 1,  
the HCG25 gene for HLA complex group 25, the VP52 gene for  
vaccinia protein sorting 52 (yeast), the RPS18 gene for ribosomal  
protein S18, the B3GALT4 gene for UDP-Gal:betaGalNAc 4-epi-  
1,3-galactosyltransferase, polypeptide 4, the 3' end of the Gcorf11  
gene for chromosome 6 open reading frame 11, and 7 CpG islands,  
complete sequence.

AL844527 109523 bp DNA linear PRI 18-MAY-2005  
HTG: B3GALT4; G6ORF11; COL11A2; CPG island; HCG25; HSD17B8;  
HTRAF1P; RING1; RPS18; RXR; SLC39A7; VP52; ZNF314P.

## SOURCE

## ORGANISM

Homo sapiens (human)  
Eukaryota; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominoidea; Homo; Homo sapiens

REFERENCE 1 (bases 1 to 109523)  
AUTHORS Leongamornlert, D.  
TITLE Direct Submissions  
JOURNAL Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk

## COMMENT

On Nov 1, 2002 this sequence version replaced gi:22204803.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information  
on the WORMPEP database can be found at  
http://www.sanger.ac.uk/projects/C\_elegans/wormpep This sequence  
was generated from part of bacterial clone contigs constructed by  
the MHC HaploType Consortium and collaborators. Further information  
can be found at  
http://www.sanger.ac.uk/HGP/Chr6/MHC  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: vegas@sanger.ac.uk  
-----

This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., piped quality >  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one subclone; and the assembly was confirmed by restriction digest,

except on the rare occasion of the clone being a YAC.  
DAOB-314F24 is from a DNA-arts OBL human bac library VECTOR:  
pBeloBac11.

## FEATURES

## source

1. 109523  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone="DAOB-314F24"  
/clone\_1db="DNA-arts-BAC-1.0BL.1"

## misc\_feature

gene  
join(complement(201..20909), .93842)

## mRNA

1  
/locus\_tag="DAOB-79P13.8-001"  
join(complement(20615..20909), complement(17776..17925),  
complement(17434..17644), complement(16818..16980),  
complement(15083..15274), complement(13448..13510),  
complement(9558..9627), complement(9421..9462),  
complement(9125..9187), complement(8715..8789),  
complement(8125..8262), complement(7888..7944),  
complement(7688..7741), complement(7530..7583),  
complement(7377..7430), complement(7136..7189),  
complement(6892..6945), complement(6765..6809),  
complement(6589..6642), complement(6119..6163),  
complement(5869..5922), complement(5638..5682),  
complement(5449..5502), complement(5178..5222),  
complement(4890..4943), complement(4716..4760),  
complement(4473..4526), complement(4031..4138),  
complement(2984..3037), complement(2622..2675),  
complement(2458..2512), complement(2230..2283),  
complement(2158..2202), complement(2001..2012),  
complement(1845446..7:93801..93842),  
complement(1845446..7:93647..93700),  
complement(1845446..7:93359..93412),  
complement(1845446..7:92829..92967),  
complement(1845446..7:92588..92677),  
complement(1845446..7:92349..92402),  
complement(1845446..7:92012..92119),  
complement(1845446..7:91766..91873),  
complement(1845446..7:91570..91663),  
complement(1845446..7:91413..91466),  
complement(1845446..7:91110..91217),  
complement(1845446..7:90857..90910),  
complement(1845446..7:90626..90679),  
complement(1845446..7:90347..90400),  
complement(1845446..7:90131..90184),  
complement(1845446..7:89683..89780),  
complement(1845446..7:89255..89308),  
complement(1845446..7:89005..89058),  
complement(1845446..7:88818..88871),  
complement(1845446..7:88098..88151),  
complement(1845446..7:87725..87823),  
complement(1845446..7:87563..87616),  
complement(1845446..7:87369..87422),  
complement(1845446..7:87019..87126),  
complement(1845446..7:86812..86865),  
complement(1845446..7:86470..86505),  
complement(1845446..7:86230..86283),  
complement(1845446..7:85848..86115),  
complement(1845446..7:85151..85263),  
complement(1845446..7:84566..84772),  
complement(1845446..7:82994..84117)  
/gene="COL11A2"  
/locus\_tag="DAOB-79P13.8-001"  
/product="collagen type XI alpha 2"  
/note="match: EST: A118676.1 BU64080.1 BX282345.1  
match: cDNAs: J04974.1 U18997.1 U6789.1"  
join(complement(20615..20656), complement(17776..17925),  
complement(17434..17644), complement(16818..16980),  
complement(15083..15274), complement(13448..13510),

## CDS

```

complement (9568, .9627), complement (9421, .9462),
complement (9125, .9187), complement (8715, .8789),
complement (8176, .8262), complement (7888, .7944),
complement (7688, .7741), complement (7530, .7589),
complement (7377, .7430), complement (7136, .7188),
complement (6892, .6945), complement (6765, .6809),
complement (6589, .6642), complement (6119, .6163),
complement (5869, .5922), complement (5638, .5682),
complement (5449, .5502), complement (5178, .5222),
complement (4890, .4943), complement (4716, .4760),
complement (4473, .4526), complement (4031, .4188),
complement (2984, .3037), complement (2622, .2675),
complement (2468, .2512), complement (2330, .2383),
complement (2158, .2202), complement (2001, .2012),
complement (AL845446, .7.93801, .93842),
complement (AL845446, .7.93647, .93700),
complement (AL845446, .7.93359, .93412),
complement (AL845446, .7.92829, .92936),
complement (AL845446, .7.92588, .92677),
complement (AL845446, .7.92349, .92402),
complement (AL845446, .7.92012, .92119),
complement (AL845446, .7.91766, .91873),
complement (AL845446, .7.91570, .91623),
complement (AL845446, .7.91413, .91466),
complement (AL845446, .7.91110, .91217),
complement (AL845446, .7.90657, .90910),
complement (AL845446, .7.90626, .90679),
complement (AL845446, .7.90347, .90400),
complement (AL845446, .7.90131, .90184),
complement (AL845446, .7.89683, .89790),
complement (AL845446, .7.89255, .89308),
complement (AL845446, .7.89005, .89058),
complement (AL845446, .7.88818, .88871),
complement (AL845446, .7.88098, .88151),
complement (AL845446, .7.87725, .87832),
complement (AL845446, .7.87563, .87616),
complement (AL845446, .7.87369, .87422),
complement (AL845446, .7.87019, .87126),
complement (AL845446, .7.86812, .86865),
complement (AL845446, .7.86470, .86505),
complement (AL845446, .7.86230, .86283),
complement (AL845446, .7.85848, .86115),
complement (AL845446, .7.85151, .85263),
complement (AL845446, .7.84566, .84772),
complement (AL845446, .7.83977, .84117),
/gene="COL11A2"
/locus_tag="DAOB-79P13.8-001"
/standard_name="OTHUMP00000014556"
/notes=match: proteins: Q28083 Q81ZC6 Q8UUM5 Q9JL12"
/codon_start=1
/product="collagen type XI alpha 2"
/protein_id="CAI41834.1"
/db_xref="GI:57209737"
/db_xref="GOA:O5JP94"
/db_xref="InterPro:IPR000885"
/db_xref="InterPro:IPR001791"
/db_xref="InterPro:IPR003129"
/db_xref="InterPro:IPR008160"
/db_xref="InterPro:IPR008161"
/db_xref="UniProt/TREMBL:O5JP94"
/translation="MERCSRCHLLLLPLVLTGLSAPAGAPVDVLRALRPSLP
DYSAGVRLGELGAPVRLVLEDPGRPSOPRPLSLADGKHRAVAIVGGS
VTLIDCKRTVRPLPRSRNPVLTDTGVTILFARLIDEVFGEDVOLAIVGVQAY
ESRQKELRBSGGQRERPNQGRSPQQRSLRHPQNERPQSDPTGEEBEI
LSSSLPLLEBAHGRGLKGRKBPRAVLEPMLVGGPGRGPAQLIGPQIQNG
PVGDPERGPGRAGLPGSDGAPGPTGLMPLFRSGCGKGPVVAAGPQAAIL
OQARLALRGPGRPMGYTGRPLGPGSGGLKGSGLDPOGRGPGQGLTGPGKAL
RERAGADGARMGDPGVKGDGPGDGLPGLPGRKGRDPTGAQGLRPGPDGGRGD
GRTGRGLPGRSGRGLLPKGRPGI RPPGRGMDGPGRPGSLGPOGRRPPQGS
TRETQGLRPGQALIGHGKGPQGRKGLGMPGSDPDPGHGPGEGPGTKGQGPBG
QEPILGPRGVRGVGDI RGLKGRKGRKEDPGRKGDIGVKGDDGEGVGSRGED
GEGPRGRTGPTGDPGRPLMGKGRKGLGVPGDLPGRGQGPGRSLRPPGRGASGKGD

```

```

ARGLKSXPRGERGPTPRGGRGRTGSKGAKTSGGDPGHPGRGLPGPGP
NGFPPKPRPGPDGLRPHGGRGRTGTPPGRPVGVGQAGAGTGMGR
GHPGRPGREGULRGTAKGRTKDDPPRGA RGDGPAGLGVPGRGLPTAGGPG
Query Match 1.7%; Score 53; DB 8; Length 109523;
Best Local Similarity 100.0%; Pred. No. 1.6e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 2888 TTAGCAGGTGATCACCCTGAGCCAGAGTTCGAGACAGGCTGCCAAT 2340
DB 66876 TTAGCAGGTGATCACCCTGAGCCAGAGTTCGAGACAGGCTGCCAAT 66928
RESULT 23
AL360268/c
LOCUS
DEFINITION
AL360268 110130 bp. DNA linear. PRI 18-MAY-2005
Human DNA sequence from clone RP11-379C10 on chromosome 9 contains
the gene for a novel protein, the 5' end of the PCSCL gene for
likely ortholog of rat peroxisomal Ca-dependent solute carrier-1like
protein (KIAA1896), two novel genes and a Cpg island, complete
sequence.
ACCESSION
AL360268
VERSION
AL360268.14 GI:19351891
KEYWORDS
HTG; Cpg Island; KIAA1896; PCSCL.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 110130)
REFERENCE
Lloyd, D.
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Mar 11, 2002 this sequence version replaced gi:18650682.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 9, constructed by the Sanger Centre Chromosome 9 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr9
RP11-379C10 is from the library RPCI-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACe3.6
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
location/Qualifiers
1..110130
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="RZPD:RPCI18753C10379"
/db_xref="Taxon:9606"
/chromosome="9"
/clone="RP11-379C10"
/clone_lib="RPCI-11.2"
1
/notes="Clone_left_end; RP11-379C10"
misc_feature

```

misc_feature	303	/note="Clone_right_end: RP11-20324"
gene		/complement(join(48570..48742,66591)..66797))
mRNA		/locus_tag="RP11-379C10.1-001"
		/complement(join(48570..48742,66591)..66797))
		/locus_tag="RP11-379C10.1-001"
		/note="match: ESTs: AA961281.1"
misc_feature	64768	
		/note="Tandem repeat. Forced join. Gap size estimated to be approximately 200bp by restriction digest data."
gene		/complement(join(70668..73335,76026)..76756)
		/gene="RP11-379C10.2"
mRNA		/locus_tag="RP11-379C10.2-001"
		/complement(join(70668..73335,76026)..76756)
		/gene="RP11-379C10.2"
		/locus_tag="RP11-379C10.2-001"
		/note="match: ESTs: AA043145.1, AA340072.1, AA359485.1, A1431105.1, AA631602.1, AA632403.1, A1081135.1, A1147614.1, A11313258.1, A18664878.1, A10682728.1, AM449929.1, AW656466.1, AM638390.1, AM956385.1, AM963898.1, BE044419.1, BE293321.1, BM385388.1, BE741976.1, BE131845.1, BF724763.1, BF724764.1, BF796375.1, BG009209.1, BC340504.1, BM016966.1, BM464251.1, BM510367.1, BM511394.1, BM678197.1, BM725183.1, BM740427.1, BM740504.1, BM740527.1, BM750316.1, BM757831.1, BM759701.1, BM853143.1, BM905218.1, BM913399.1, BM931058.1, BQ231034.1, BQ441346.1, BU059062.1, BU165007.1, CA310459.1, CA312400.1, M69444.1
		BUL65007.1, CA310459.1, CA312400.1, M69444.1
		match: CDNA: AB060237.1, AE833392.1"
		/complement(join(70668..71261,72809..73177))
		/gene="RP11-379C10.2"
gene		/locus_tag="RP11-379C10.2-004"
		/complement(join(70668..71261,72809..73177))
		/gene="RP11-379C10.2"
mRNA		/locus_tag="RP11-379C10.2-004"
		/note="match: ESTs: AA281868.1, BG008293.1, BG912959.1, BM684531.1, BQ370637.1"
		/complement(70668)
		/gene="RP11-379C10.2"
polyA_site		/locus_tag="RP11-379C10.2-004"
		/complement(70672)
		/gene="RP11-379C10.2"
polyA_site		/locus_tag="RP11-379C10.2-004"
		/complement(70674)
		/gene="RP11-379C10.2"
polyA_signal		/locus_tag="RP11-379C10.2-004"
		/complement(70691..70696)
		/gene="RP11-379C10.2"
gene		/locus_tag="RP11-379C10.2-004"
		/complement(join(71425..73335,75582..75719))
		/gene="RP11-379C10.2"
mRNA		/locus_tag="RP11-379C10.2-002"
		/complement(join(71425..73335,75582..75719))
		/gene="RP11-379C10.2"
		/locus_tag="RP11-379C10.2-002"
		/note="match: ESTs: BE747080.1
		match: CDNAs: BC013371.1"
		/complement(72723)
		/gene="RP11-379C10.2"
polyA_site		/locus_tag="RP11-379C10.2-004"
		/complement(72731)
		/gene="RP11-379C10.2"
polyA_signal		/locus_tag="RP11-379C10.2-004"
		/complement(72743..72748)
		/gene="RP11-379C10.2"
CDS		/locus_tag="RP11-379C10.2-004"
		/complement(join(72863..73335,76026..76536))
		/locus_tag="RP11-379C10.2-001"
		/standard_name="OTTHMP00000022227"
		/note="match: proteins: Q8WU12 Q9BE21"
		/codon_start=1
		/protein_id="CAH73132.1"
		/db_xref="GI:55664183"

[illegible]

RESULT 24	AP000247	113919 bp	DNA	linear	PRI 15-MAR-2003
AP000247	Homo sapiens genomic DNA, chromosome 21q, clone:RP5-871G18,				
LOCUS	complete sequence.				
DEFINITION	Homo sapiens genomic DNA, chromosome 21q, clone:RP5-871G18,				
ACCESSION	AP000247				
VERSION	AP000247.2				
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Hattori, M., Ishi, K., Toyoda, A., Taylor, T.D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.				
TITLE	Homo sapiens genomic DNA				
JOURNAL	Published Only in Database (1999)				
REFERENCE	2 (bases 1 to 113919)				
AUTHORS	Hattori, M., Ishi, K., Toyoda, A., Taylor, T.D., Hong-Seog, P., Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.				
TITLE	Direct Submission				
JOURNAL	Submitted (13-MAY-1999) Masahira Hattori, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Saito-cho, Tsukuba, Ibaraki, Japan, 305-8565, Japan (E-mail: hattori@gsc.riken.go.jp, URL: http://hgp-gsc.riken.go.jp/, Tel:81-45-503-9111, Fax:81-45-503-9170)				
COMMENT	On Jul 17, 2001 this sequence version replaced gi:4835616.				
FEATURES	Location/Qualifiers				
SOURCE	1..113919				
ORIGIN	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	/chromosome="21"				
	/map="21q"				
	/clone="RP5-871G18"				
Query Match	1.7%; Score 53; DB 8; Length 113919;				
Best Local Similarity	100.0%; Pident. No. 1.0e-16;				
Matches	53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				
Qy	2888 TGAGGCGAGTGTGATCAGCTGAGGCGAGGATTCGAGACCAAGCTGGCCAAAT 2940				
Db	13942 TGAGGCGAGTGTGATCAGCTGAGGCGAGGATTCGAGACCAAGCTGGCCAAAT 13994				
RESULT 25	HSJ733M16				
LOCUS	120825 bp				
DEFINITION	Human DNA sequence from clone RP4-733M16 on chromosome 1p36.11-36.23 contains four novel genes, a novel gene (PLUJ3962), a metalloprotein family pseudogene and three CpG islands, complete sequence.				
ACCESSION	AL109627.18				
VERSION	AL109627.18				
KEYWORDS	HTG, PLUJ3962, KIAA1332, MGC10731.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.				
AUTHORS	Hall, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,				

```
COMMENT
CambridgeShire, CB10 1SA, UK; E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequest@sanger.ac.uk
On Jan 24, 2000 this sequence version replaced gt:6723665.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TREMBL; Wp:, WormPep; Information on the WormPep database can be found at http://www.sanger.ac.uk/projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chrl RP4-733M16 is from the library RPCI-4 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm VECTOR: pCYPAC2
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.
Location/Qualifiers
FEATURES
source
1..120825
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="RZPD:RPCIP704M16733"
    /db_xref="taxon:9606"
    /chromosome="1"
    /map="p36.11-36.23"
    /clone="RP4-733M16"
    /clone_1id="RPCI-4"
1
    /note="Clone left end: RP4-733M16"
    join(complement(ALJ558794.19:480..678),10770..10886,21772..24375)
        /gene="FBXO42"
        /locus_tag="RP4-733M16.6-002"
        join(complement(ALT558794.19:480..678),10770..10886,21772..24375)
            /gene="FBXO42"
            /locus_tag="RP4-733M16.6-002"
            /product="F-box protein 42"
            /note="match: cDNAs: Em:AL833207.1"
33887
    /note="Possible local misassembly of alu containing reads."
Overall assembly confirmed by restriction digest."
33888
    /note="Possible local misassembly of alu containing reads."
Overall assembly confirmed by restriction digest."
33892
    /note="Possible local misassembly of alu containing reads."
Overall assembly confirmed by restriction digest."
33903
    /note="Possible local misassembly of alu containing reads."
Overall assembly confirmed by restriction digest."
33910
    /note="Possible local misassembly of alu containing reads."
Overall assembly confirmed by restriction digest."
33921
    /note="Possible local misassembly of alu containing reads."
Overall assembly confirmed by restriction digest."
33933
    /note="Possible local misassembly of alu containing reads."
Overall assembly confirmed by restriction digest."
33938
    /note="Possible local misassembly of alu containing reads."
Overall assembly confirmed by restriction digest."
```

Query Match	1.7%	Score 53	DB 8	Length 120825
Best Local Similarity	100.0%	Pred. No. 1.6e-16		
Matches 53	Conservative 0	Mismatches 0	Indels 0	Gaps 0
Db	2888	TGAGCGAGTGTGATTCACCTGAGGCGAGAGGTTGAGAGCGAGCTGGCCACAT	2940	
	52626	TGAGCGAGTGTGATTCACCTGAGGCGAGAGTTCGAGACCGAGCTGGCCACAT	52678	
RESULT 26	AC095044/c	134184 bp	DNA	linear PRI 15-OCT-2003
LOCUS	AC095044			
DEFINITION	Homo sapiens BAC clone RP11-132B16 from 7, complete sequence.			
ACCESSION	AC095044	AC023954		
VERSION	AC095044.3	GI:18464270		
KEYWORDS	HTG.			
ORGANISM	Homo sapiens (human)			
SOURCE	Homo sapiens			
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
REFERENCE	Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;			
AUTHORS	Hominidae; Homo			
TITLE	1 (bases 1 to 134184)			
JOURNAL	Sulston, J.E. and Wilson, R.			
PUBMED	Toward a complete human genome sequence			
REFERENCE	Genome Res. 8 (11), 1097-1108 (1998)			
AUTHORS	9847074			
TITLE	2 (bases 1 to 134184)			
JOURNAL	Cotton, M. and VanBrunt, A.			
PUBMED	The sequence of Homo sapiens BAC clone RP11-132B16			
REFERENCE	Unpublished (2001)			
AUTHORS	3 (bases 1 to 134184)			
TITLE	Waterston, R.H.			
JOURNAL	Direct Submission			
PUBMED	Submitted (15-SEP-2001) Genome Sequencing Center, Washington			
REFERENCE	University School of Medicine, 4444 Forest Park Parkway, St. Louis,			
AUTHORS	MO 63108, USA			

REFERENCE 4 (bases 1 to 134184)  
AUTHORS Waterston,R.H.  
TITLE Direct Substition  
JOURNAL Submitted (01-FEB-2002) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
5 (bases 1 to 134184)  
AUTHORS Wilson,R.  
TITLE Direct Substition  
JOURNAL Submitted (15-OCT-2003) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Feb 1, 2002 this sequence version replaced gi:16418233.  
----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu>  
Contact: [saplens@wuston.wustl.edu](mailto:saplens@wuston.wustl.edu)  
----- Summary Statistics  
Center Project name: H\_NH0132B16  
Drafting Center: WIBR  
-----  
NOTICE: This sequence may not represent the entire insert of this  
clone. It may be shorter because we only sequence overlapping  
clone sections once, or longer because we provide a small overlap  
between neighboring data submissions.  
  
This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate  
chemistry, or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by sequence  
from more than one subclone; and the assembly was confirmed by  
restriction digest.  
  
MAPPING INFORMATION:  
The sequence of this clone was established as part of a mapping and  
sequencing collaboration between the NHGRI Chromosome 7 Mapping  
Project (Eric D. Green, Director), John D. McPherson in the  
Department of Genetics (Washington University), and the Washington  
University Genome Sequencing Center. For additional information  
about the map position of this sequence, see  
<http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send  
<mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu>  
  
SOURCE INFORMATION:  
The RPCT-11 human BAC library was made from the blood of one male  
donor, as described by Osoegawa,K., Moon,P.Y., Zhao,B., Frengen,E.,  
Tateno,M., Cataneese,J.J. and de Jong,P.J. (1998) An improved  
approach for construction of bacterial artificial chromosome  
libraries. Genomics 51:1-8. The clone may be obtained either from  
Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong  
and coworkers at <http://www.chori.org>  
VECTOR: pBACe3.6.  
Location/Qualifiers  
1..134184  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="7"  
/map="7"  
/clone="RP11-132B16"  
/clone\_1fb="RPCT-11"  
381..697  
/rpt\_family="Alu"  
712..1032  
/rpt\_family="Alu"  
1033..1060  
/rpt\_family="AT\_rich"  
1276..1592  
/rpt\_family="MalR"  
1594..1772  
/rpt\_family="(TTCC)n"

```
repeat_region 1775..1852 /rpt_family="CT-rich"
repeat_region 1853..1985 /rpt_family="L2"
repeat_region 2062..2086 /rpt_family="Alu"
repeat_region 2087..2109 /rpt_family="Alu"
repeat_region 2110..2260 /rpt_family="(TTAA)n"
repeat_region 2261..2565 /rpt_family="Alu"
repeat_region 2566..2704 /rpt_family="Alu"
repeat_region 2858..2887 /rpt_family="Alu"
repeat_region 3064..3375 /rpt_family="(TTCA)n"
repeat_region 3381..3471 /rpt_family="ERV1"
repeat_region 3600..3991 /rpt_family="ERV1"
repeat_region 4326..4504 /rpt_family="L2"
repeat_region 4517..4597 /rpt_family="(TGA)n"
repeat_region 4613..4996 /rpt_family="(TGA)n"
repeat_region 5710..5779 /rpt_family="L1"
repeat_region 6510..6573 /rpt_family="MIR"
repeat_region 6634..6937 /rpt_family="(GGA)n"
repeat_region 7031..7083 /rpt_family="Alu"
repeat_region 7095..7621 /rpt_family="A-rich"
repeat_region 7436..7642 /rpt_family="ERV1"
repeat_region 7653..8019 /rpt_family="ERV1"
repeat_region 8223..8536 /rpt_family="MalR"
repeat_region 8599..8876 /rpt_family="Alu"
repeat_region 8877..8903 /rpt_family="Alu"
repeat_region 9039..9092 /rpt_family="(TTAA)n"
repeat_region 9093..9259 /rpt_family="(TTCC)n"
repeat_region 9260..9556 /rpt_family="(TTTC)n"
repeat_region 10048..10363 /rpt_family="Alu"
repeat_region 10364..10394 /rpt_family="Alu"
repeat_region 10566..10725 /rpt_family="(GA)n"
repeat_region 10725..10908 /rpt_family="ERV1"
repeat_region 10905..11207 /rpt_family="Alu"
repeat_region 11209..11470 /rpt_family="Alu"
repeat_region 11638..12115 /rpt_family="ERV1"
repeat_region 12182..12491 /rpt_family="MER1_type"
repeat_region 12519..12696 /rpt_family="Alu"
repeat_region 13006..13037 /rpt_family="MIR"
```



```

repeat_region      /rpt_family="AT_rich"
                    13035..13317
repeat_region      /rpt_family="Alu"
                    13680..13793
repeat_region      /rpt_family="MIR"
                    13814..14096
repeat_region      /rpt_family="Alu"
                    14185..15007
repeat_region      /rpt_family="L1"
                    15050..15128
repeat_region      /rpt_family="ERVU"
                    15163..15327
repeat_region      /rpt_family="ERVU"
                    16005..16166
repeat_region      /rpt_family="Achoho"
                    16167..16471
repeat_region      /rpt_family="Alu"
                    16564..16913
repeat_region      /rpt_family="Alu"
                    17764..17911
repeat_region      /rpt_family="Alu"
                    17926..18141
repeat_region      /rpt_family="L1"
                    18161..18681
repeat_region      /rpt_family="MALR"
                    18902..18984
repeat_region      /rpt_family="GA-rich"
                    19076..19203
repeat_region      /rpt_family="L1"

```

Query Match 1.7%; Score 53; DB 8; Length 134184;  
Best Local Similarity 100.0%; Pred. No. 1.6e-16;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2888 TGAGGAGGTGATCCTGAGGCCGAGGTTGAGACCCAGCTGCGCAACAT 2940  
Db 37245 TGAGGAGGTGATCCTGAGGCCGAGGTTGAGACCCAGCTGCGCAACAT 37193

RESULT 27 BS000110 157686 bp DNA linear PRI 12-JUN-2004  
LOCUS BS000110  
DEFINITION Pan troglodytes chromosome 22 clone:PTB-090P24, map 22, complete  
sequences.

ACCESSION BS000110 BA000045  
VERSION BS000110.1 GI:37537377  
KEYWORDS HTG.  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Pan.

REFERENCE 1  
AUTHORS The International Chimpanzee Chromosome 22 Consortium.  
TITLE DNA sequence and comparative analysis of chimpanzee chromosome 22  
JOURNAL Nature 429, 382-388 (2004)  
REFERENCE 2 (bases 1 to 157686)  
AUTHORS Hattori,M., Toyoda,A., Watanabe,H., Taylor,T.D., Kuroki,Y.,  
Fujiyama,A. and Sakaki,Y.  
TITLE Direct Submision  
JOURNAL Submitted (12-MAY-2003) Masahira Hattori, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),  
1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail:hattori@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/  
Tel:81-45-503-9111, Fax:81-45-503-9170)  
The Chimpanzee Chromosome 22 Sequencing Consortium consists of:  
\*Chinese National Human Genome Center at Shanghai, Shanghai, China;  
\*GBF, Dept. of Genome Analysis, Braunschweig, Germany; \*Institute  
of Molecular Biotechnology, Jena, Germany; \*KIRB Genome Research  
Center, Daejeon, Korea;  
\*Max-Planck-Institute for Molecular Genetics, Berlin, Germany;  
\*National Institute of Genetics, Mishima, Japan;  
\*National Yang Ming University Genome Research Center, Taipei,

Taiwan;  
\*RIKEN Genomic Sciences Center, Yokohama, Japan.

----- Genome Center  
Center: RIKEN Genomic Sciences Center  
Center code: RIKEN  
Web site: http://hgp.gsc.riken.go.jp/  
Contact: hattori@gsc.riken.go.jp  
----- Project Information  
Center project name: The Chimpanzee Chromosome 22 Sequencing Project  
Center clone name: PTB-090P24  
----- Summary Statistics  
Sequencing vector: pUC18,pUC13,PT219R; 100% of reads Chemistry:  
Dye-terminator Big Dye and ET; 100% of reads Assembly program:  
Phrap; version 0.990329

-----  
This sequence was finished as follows unless otherwise noted: all  
regions were double stranded, sequenced with an alternate  
chemistry, or covered by high quality data (i.e., phred quality >= 30);  
an attempt was made to resolve all sequencing problems, such as  
compressions and repeats; all regions were covered by at one  
subclone or more than one M13 subclone;  
and the assembly was confirmed by restriction digest.

#### Source information:

The RPI-43 chimpanzee BAC library was prepared from DNA isolated  
from the blood of a single male chimpanzee using published  
protocols (Osoegawa,K. et al. Genomics 52:1-8). The DNA from the  
chimpanzee ('Clint') was obtained from the Yerkes Primate Center in  
Atlanta. The library was prepared by BaoLi Zhu, Chung Li Shu,  
Kazutoyo Osoegawa, Evan Eichler & Pieter J de Jong. The library  
characteristics are described at  
http://www.chori.org/bacpac/mchimp43.htm.  
The clone may be obtained from Pieter J. de Jong and coworkers  
(http://www.chori.org/bacpac).  
VECTOR: pBACe3.6

The CHORI-251 chimpanzee BAC library was prepared from DNA isolated  
from the blood of a single male chimpanzee using published  
protocols (Osoegawa,K. et al. Genomics 52:1-8). The DNA from the  
chimpanzee ('Clint') was obtained from the Yerkes Primate Center in  
Atlanta. The library was prepared by BaoLi Zhu, Chung Li Shu,  
Kazutoyo Osoegawa, Evan Eichler & Pieter J de Jong. The library  
characteristics are described at  
http://www.chori.org/bacpac/chimpanzee251.htm.  
The clone may be obtained from Pieter J. de Jong and coworkers  
(http://www.chori.org/bacpac).

VECTOR: PTARBAC2.1  
The PTB1 chimpanzee BAC library was prepared from DNA isolated from  
cultured cells established from the blood of a single male  
chimpanzee.

Clones may be obtained from Aaao Fujiyama and co-workers  
(http://www.gsc.riken.go.jp).  
VECTOR: pKS145

The PT22 chimpanzee Fosmid library was prepared from DNA isolated  
from cultured cells established from the blood of a single male  
chimpanzee.  
Clones may be obtained from Aaao Fujiyama and co-workers  
(http://www.gsc.riken.go.jp).  
VECTOR: pKS143

#### Sequence Quality Assessment:

This entry has been annotated with sequence  
estimates computed by the Phrap assembly program.  
All manually edited bases have been reduced to quality zero.  
Quality levels above 40 are expected to have less than 1 error in  
10,000 bp.

-----  
Neighboring clones: RP43-150L08(left) and PTB-093X04(right).

#### FEATURES

source  
location/Qualifiers  
1..157686  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"

```

/db xref="taxon:9598"
/chromosome="22"
/clone="PTB-090P24"
/clone_lib="PTB1 chimpanzee BAC"

ORIGIN
    Query Match      1.7%; Score 53; DB 8; Length 157686;
    Best Local Similarity 100.0%; Pred. No. 1,6e-16;
    Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2888 TGAGGAGGTGATGATCCTGAGGCCAGAGTTGAGACCAAGCTGGCCACAT 2940
        |||
DB      123811 TGAGGAGGTGATCCTGAGGCCAGAGTTGAGACCAAGCTGGCCACAT 123863

RESULT 28
AL645940      158033 bp      DNA      linear      PRI 18-MAY-2005
LOCUS      Human DNA sequence from clone Xxbac-157A10 on chromosome 6 contains
DEFINITION  the 3' end of the gene for a novel protein similar to major
             histocompatibility complex, class II, DP beta 2 pseudogene, a major
             histocompatibility complex, class II, DP beta 1 pseudogene, two
             putative novel transcripts, the COL11A2 gene for collagen, type XI,
             alpha 2, the RXRB gene for retinoid X receptor, beta, the HKE4 gene
             for HLA-class II region expressed gene KE4, the HKE4 gene for
             hydroxysteroid (17-beta) dehydrogenase 8, the RING1 gene for ring
             finger protein 1, a zinger finger protein pseudogene, a HIV TAT
             specific factor 1 (HTRSF1) pseudogene, a novel transcript, the
             SACC2L gene for SACC2 suppressor of actin mutations 2-like (yeast),
             the 5' end of the RPS18 gene for ribosomal protein S18 and six Cpg
             islands, complete sequence.
AL645940      158033 bp      DNA      linear      PRI 18-MAY-2005
LOCUS      HTG; ABR-1, ABR1; class II region expressed gene KE4; COL11A2;
DEFINITION  collagen, type XI, alpha 2; Cpg island; D6S115B; D6S2244B;
             D6S2245B; DFN13; FABG; FABG1; H-2RIIB; H2-KE4; H2-KE6; HIV TAT
             specific factor 1; HKE3; HKE4; HKE5; HKE6; HLA-DPB3; HSD17B8;
             HTRSF1; hydroxysteroid dehydrogenase 8; KE-3; KE3; KE4; KE6; major
             histocompatibility complex, class II; major histocompatibility
             complex, class II, DP bet; NR2B2; PARP; RCor-1; retinoid X
             receptor, beta; ribosomal protein S18; ring finger protein 1;
             RING1; RING2; RING5; RNFL; RPS-103B10; RPS18; RXRB; SACC2; SACC2
             suppressor of actin mutations 2-like; SACC2L; STL3; TAT-SF1; zinc
             finger protein.
Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Homiidae; Homo.
REFERENCE   1 (bases 1 to 158033)
AUTHORS    Leongamornlert,D.
TITLE      Direct SubMISSION
JOURNAL     Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
             Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
             Clone requests: clonerequest@sanger.ac.uk
             On Feb 6, 2002 this sequence version replaced gi:18070925.
             The following abbreviations are used to associate primary accession
             numbers given in the feature table with their source databases:
             Em; EMBL; Swi; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information
             on the WORMPEP database can be found at
             http://www.sanger.ac.uk/Projects/C_elegans/wormpep/xbac-157A10
             is from a CHORI-501 human bac - PEP cell line library VECTOR:
             PTABAC2.1
             This sequence was generated from part of bacterial clone contigs
             constructed by the MHC Haplotype Consortium and collaborators.
             Further information can be found at
             http://www.sanger.ac.uk/HGP/Chr6/MHC
             -----Genome Center
             Center: Wellcome Trust Sanger Institute
             Center code: SC
             Web site: http://www.sanger.ac.uk
             Contact: vegas@sanger.ac.uk
             -----

```

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

```

FEATURES
    source
        1..158033
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="6"
            /clone_lib="CHORI-501"
            /clone_1ib="Xxbac-157A10"
            /clone_1ib="CHORI-501"
            /note="Clone right end: Xxbac-138A21"
            join(AL645931.7:120452..120551,AL645931.7:124893..124899,
            2001..2257,12865..13125,13671..13793,14102..14143)
            /locus_tag="Xxbac-BPG157A10.5-001"
            /pseudo
            join(AL645931.7:120452..120551,AL645931.7:124893..124899,
            2001..2257,12865..13125,13671..13793,14102..14143)
            /locus_tag="Xxbac-BPG157A10.5-001"
            /note="match: proteins: 019702 P01916 P13763 Q30166 Q31163
            Q31633 Q95HC1 Q95IH2"
            /pseudo
            /codon_start=1
            /product="major histocompatibility complex, class II, DP
            beta 2 (pseudogene)"
            join(AL645931.7:120387..120551,AL645931.7:124893..124899,
            2001..2257,3428..4186)
            /locus_tag="Xxbac-BPG157A10.4-001"
            join(AL645931.7:120387..120551,AL645931.7:124893..124899,
            2001..2257,3428..4186)
            /locus_tag="Xxbac-BPG157A10.4-001"
            /note="match: ESTs: A1971689.1 BG534679.1
            match: cDNAs: BC017967.1"
            complement(16239..16389)
            /locus_tag="Xxbac-BPG157A10.3-001"
            complement(16239..16389)
            /pseudo
            complement(16239..16389)
            /locus_tag="Xxbac-BPG157A10.3-001"
            /note="match: proteins: P20036 Q8SP47 Q9TPX3 Q9XS10"
            /pseudo
            /codon_start=1
            /product="major histocompatibility complex, class II, DP
            alpha 2 (pseudogene)"
            join(29820..29983,32252..32789)
            /locus_tag="Xxbac-BPG157A10.15-001"
            join(29820..29983,32252..32789)
            /locus_tag="Xxbac-BPG157A10.15-001"
            /note="match: ESTs: BG033647.1"
            complement(47731..77536)
            /gene="COL11A2"
            /locus_tag="Xxbac-BPG157A10.2-001"
            complement(join(47731..48129,48779..48855,49304..49510,
            49889..50001,50586..50853,50968..51021,51208..51243,
            51550..51603,51757..51864,52107..52160,52301..52354,
            52463..52570,52836..52889,53556..53609,53743..53796,
            53929..54045,54420..54527,54668..54684,55137..55137,
            55363..55416,55594..55647,55954,56151..56204,
            56308..56361,56504..56611,56750..56857,57087..57140,
            57326..57415,57567..57674,58097..58150,58385..58438,
            58539..58592,58738..58782,59048..59092,59106,
            59202..59255,59564..59617,60611..60718,61053..61106,
            61296..61340,61470..61523,61758..61802,62029..62082,
            62218..62262,62449..62502,62699..62743,63169..63222,
            63345..63389,63472..63525,63716..63769,63957..64010,
            64110..64163,64268..64321,64468..64524,64756..64842,
            65295..65369,65705..65767,66001..66042,66148..66207,
            70029..70091,71664..71855,73399..73561,74015..74225,
            74357..74506,77196..77536))

```



Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B., Choquel,Y., Colangelo,M., Collins,S., Collimore,A., Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Kartae,A., Kelle,C., Larocque,K., Lamazares,R., Landers,I., Lehoczy,J., Levine,R., Liu,G., Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., Meldrum,J., Menues,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunthang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Scojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zaitoun,J., Zembek,L., Zimmer,A. and Zody,M.

Direct Submission  
Submitted (20-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Feb 20, 2002 this sequence version replaced gi:18693515.  
All repeats were identified using RepeatMasker:  
Smit,A.F.A. & Green,P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: MIBR  
Web site: http://www-seg.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L6420  
Center clone name: 210\_K\_20  
-----

FEATURES  
source  
-----  
Location/Qualifiers  
1..171051  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="18"  
/map="18"  
/clone="RP11-210K20"  
/clone\_1lb="RPC1-11 Human Male BAC"  
1..176  
/rpt\_family="MIR"  
repeat\_region  
223..251  
/rpt\_family="(TG)n"  
repeat\_region  
537..643  
/rpt\_family="MER5A"  
repeat\_region  
1820..2239  
/rpt\_family="L1MA7"  
repeat\_region  
2240..2530  
/rpt\_family="AluUb"  
repeat\_region  
2531..4997  
/rpt\_family="L1MA7"  
repeat\_region  
4998..5294  
/rpt\_family="AluVa5"  
repeat\_region  
5295..5770  
/rpt\_family="L1MA7"  
repeat\_region  
complement(5840..5966)  
/rpt\_family="FLAM\_C"  
7146..7178  
/rpt\_family="AT\_rich"  
repeat\_region  
7335..7362  
/rpt\_family="(TTTA)n"  
repeat\_region  
complement(7365..7646)  
/rpt\_family="AluBo"  
repeat\_region  
complement(7678..7966)  
/rpt\_family="AluSx"  
8635..9181  
repeat\_region

/rpt\_family="L1R10A"  
complement(9652..9786)  
/rpt\_family="MER5B"  
repeat\_region  
complement(9842..10444)  
/rpt\_family="L1R9B"  
repeat\_region  
complement(10448..10690)  
/rpt\_family="HURS-P3"  
repeat\_region  
10727..11026  
/rpt\_family="AluSx"  
complement(11319..11430)  
/rpt\_family="L1R9B-int"  
repeat\_region  
complement(11427..12609)  
/rpt\_family="L1R9B-int"  
repeat\_region  
12610..12919  
/rpt\_family="AluSp"  
complement(12920..13216)  
/rpt\_family="L1R9B-int"  
repeat\_region  
13217..13527  
/rpt\_family="AluSg"  
complement(13528..13699)  
/rpt\_family="L1R9B-int"  
repeat\_region  
complement(13972..14575)  
/rpt\_family="L1R9B"  
repeat\_region  
14583..15192  
/rpt\_family="MER4B"  
complement(15330..15674)  
/rpt\_family="MER21B"  
repeat\_region  
complement(15675..15989)  
/rpt\_family="AluBo"  
complement(15990..16412)  
/rpt\_family="MER21B"  
repeat\_region  
complement(16412..16531)  
/rpt\_family="MER21B"  
repeat\_region  
complement(17688..19067)  
/rpt\_family="L1P"  
repeat\_region  
19061..19138  
/rpt\_family="L1PA13"  
repeat\_region  
19139..19201  
/rpt\_family="L1PA13"  
repeat\_region  
19202..19483  
/rpt\_family="AluSg"  
repeat\_region  
19484..19882  
/rpt\_family="L1PA13"  
repeat\_region  
21669..22281  
/rpt\_family="L1MA8"  
complement(22282..22696)  
/rpt\_family="T1gser2a"  
repeat\_region  
22697..22886  
/rpt\_family="L1MA8"  
repeat\_region  
22888..22908  
/rpt\_family="AT\_rich"  
complement(23103..23255)  
/rpt\_family="MER110"  
repeat\_region  
23290..23427  
complement(23697..24076)  
/rpt\_family="MSTC"  
repeat\_region  
complement(25025..25187)  
/rpt\_family="MIR"  
repeat\_region  
complement(25208..25502)  
/rpt\_family="AluSg"  
repeat\_region  
25814..25847  
/rpt\_family="(TTTA)n"  
complement(26782..27105)  
repeat\_region

Query Match 1.7%; Score 53; DB 8; Length 171051;  
Best Local Similarity 100.0%; Pred.No. 1.6e-16;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2888 TGAGCAGGTGATCACTGAGGCCGAGGTTTCGAGACCAAGCTTGGCCAACT 2940  
DB 90385 TGAGCAGGTGATCACTGAGGCCGAGGTTTCGAGACCAAGCTTGGCCAACT 90437



gene  
HGLQCPEDVVLGVTGTTGTTMLVPGAGEPNFDLIESNPNRKROROWEYKALJEK  
VPAELICDPPALADVDVLSLEQKKEQLEKRGYPOPAKAPOPKRGKGRSTASLV  
KRKRVMDEEHDXKROSIOQOHKHEAKRPGAPSAIDRVR"  
join(297184.1:33855..33989,297184.1:34186..34285,  
5240..5799,6368..6463,6617..6730,6853..6912)  
/gene="C6orf11"  
/locus\_tag="RP5-1033B10.11-002"  
/locus\_tag="RP5-1033B10.11-002"  
join(297184.1:33855..33989,297184.1:34186..34285,  
5240..5799,6368..6463,6617..6730,6853..6912)  
/gene="C6orf11"  
/locus\_tag="RP5-1033B10.11-002"  
/product="chromosome 6 open reading frame 11"  
/note="match: ESTs: BE164701.1 BE667540.1 BG697282.1  
BG12913.1"  
join(297184.1:31873..31978,297184.1:32067..32188,  
297184.1:33575..33725,297184.1:33854..33989,  
297184.1:34186..34285,5240..5519)  
/gene="C6orf11"  
/locus\_tag="RP5-1033B10.11-004"  
/locus\_tag="RP5-1033B10.11-004"  
join(297184.1:31873..31978,297184.1:32067..32188,  
297184.1:33575..33725,297184.1:33854..33989,  
297184.1:34186..34285,5240..5519)  
/gene="C6orf11"  
/locus\_tag="RP5-1033B10.11-004"  
/product="chromosome 6 open reading frame 11"  
/note="match: ESTs: BG470714.1"  
join(5316..5412,5705..5799,6368..6463,6617..6730,  
6853..7119)  
/gene="C6orf11"  
/locus\_tag="RP5-1033B10.11-007"  
/locus\_tag="RP5-1033B10.11-007"  
join(5316..5412,5705..5799,6368..6463,6617..6730,  
6853..7119)  
/gene="C6orf11"  
/locus\_tag="RP5-1033B10.11-007"  
/product="chromosome 6 open reading frame 11"  
/note="match: ESTs: BP476741.1"  
join(5374..5553,5705..5799,6368..6463,6617..7122)  
/gene="C6orf11"  
/locus\_tag="RP5-1033B10.11-006"  
/locus\_tag="RP5-1033B10.11-006"  
join(5374..5553,5705..5799,6368..6463,6617..7122)  
/gene="C6orf11"  
/locus\_tag="RP5-1033B10.11-006"  
/product="chromosome 6 open reading frame 11"  
/note="match: ESTs: AA599154.1 AA627958.1 A1204504.1  
A1524223.1 AM193506.1 AM294189.1 BRS04644.1 BE617532.1  
H07040.1 R3981.1 R40596.1"  
7094..7099  
/gene="C6orf11"  
/locus\_tag="RP5-1033B10.11-006"  
7111  
/gene="C6orf11"  
/locus\_tag="RP5-1033B10.11-006"  
7119  
/gene="C6orf11"  
/locus\_tag="RP5-1033B10.11-006"  
7122  
/gene="C6orf11"

polyA\_site  
polyA\_site  
polyA\_site  
polyA\_site  
polyA\_site

Query Match 1.7% Score 53; DB 8; Length 175737;  
Best Local Similarity 100.0%; Pred.No. 1.6e-16;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 2888 TGAGGAGGTGATCACTGAGGCGAGGCTCGAGACGCTGCGCAACAT 2940  
DB 47829 TGAGGAGGTGATCACTGAGGCGAGGCTCGAGACGCTGCGCAACAT 47777

RESULT 31  
AC021102  
LOCUS Homo sapiens chromosome 17 clone RP11-129G15, WORKING DRAFT  
DEFINITION  
SEQUENCE, 25 unordered pieces.  
AC021102  
ACCESSION

VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

AC021102.3 GI:8567961  
HTG, HTGS PHASE1, HTGS\_DRAFT.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 179462)  
Waterston, R.H.  
The sequence of Homo sapiens clone  
Unpublished  
2 (bases 1 to 179462)  
Waterston, R.H.  
Direct Submission  
Submitted (14-JAN-2000) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
On Jun 15, 2000 this sequence version replaced gi:7230893.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu/gsc/index.shtml>  
----- Project Information -----  
Center project name: H NH0129G15  
----- Summary Statistics -----  
Sequencing vector: M13; 75%  
Sequencing method: plasmid; 25%  
Chemistry: Dye-terminator Big Dye; 25% of reads  
Chemistry: Dye-terminator Big Dye; 25% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 165207 bases at least Q40  
Consensus quality: 169791 bases at least Q30  
Insert size: 177000; agarose-fp  
Insert size: 177062; sum-of-contigs  
Quality coverage: 3.62 in Q20 bases; agarose-fp  
Quality coverage: 3.66 in Q20 bases; sum-of-contigs  
-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 25 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1  
1892: contig of 1892 bp in length  
1893  
1992: gap of unknown length  
1993  
4443: contig of 2351 bp in length  
4444  
4443: gap of unknown length  
4444  
6559: contig of 2116 bp in length  
6560  
6559: gap of unknown length  
6560  
10517: contig of 3858 bp in length  
10518  
10617: gap of unknown length  
10618  
14073: contig of 3456 bp in length  
14074  
14173: gap of unknown length  
14174  
16899: contig of 2726 bp in length  
16900  
16999: gap of unknown length  
16900  
20655: contig of 3656 bp in length  
20656  
20755: gap of unknown length  
20756  
24533: contig of 3778 bp in length  
24534  
24633: gap of unknown length  
24634  
27750: contig of 3117 bp in length  
27751  
27850: gap of unknown length  
27851  
29522: contig of 1672 bp in length  
29523  
29622: gap of unknown length  
29623  
33750: contig of 4128 bp in length  
33751  
33850: gap of unknown length  
33851  
37904: contig of 4054 bp in length  
37905  
38004: gap of unknown length  
38005  
43341: contig of 5237 bp in length  
43342  
43341: gap of unknown length



```
* 43342 50634: contig of 7293 bp in length
* 50635 50734: gap of unknown length
* 50735 59411: contig of 8677 bp in length
* 59412 59511: gap of unknown length
* 59512 68503: contig of 8992 bp in length
* 68504 79177: contig of 10574 bp in length
* 79178 79277: gap of unknown length
* 79278 86283: contig of 7006 bp in length
* 86284 96915: contig of 10532 bp in length
* 96916 97015: gap of unknown length
* 97016 107357: contig of 10342 bp in length
* 107358 107457: gap of unknown length
* 107458 119546: contig of 12089 bp in length
* 119547 119646: gap of unknown length
* 119647 134223: contig of 14577 bp in length
* 134224 134323: gap of unknown length
* 134324 147018: contig of 12695 bp in length
* 147019 147118: gap of unknown length
* 147119 160737: contig of 13619 bp in length
* 160738 160837: gap of unknown length
* 160838 179462: contig of 18625 bp in length.
FEATURES
source
1. 179462
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/clone="RP11-129G15"
1. 1892
/note="assembly_name:Contig11"
1893. 1992
/estimated_length=unknown
1993. 4343
/note="assembly_name:Contig12"
4344. 4443
/estimated_length=unknown
4444. 6559
/note="assembly_name:Contig13"
6560. 6659
/estimated_length=unknown
6660. 10517
/note="assembly_name:Contig14"
10518. 10617
/estimated_length=unknown
10618. 14073
/note="assembly_name:Contig15"
14074. 14173
/estimated_length=unknown
14174. 16899
/note="assembly_name:Contig16"
16900. 16999
/estimated_length=unknown
17000. 20655
/note="assembly_name:Contig17"
20656. 20755
/estimated_length=unknown
20756. 24533
/note="assembly_name:Contig18"
24534. 24633
/estimated_length=unknown
24634. 27750
/note="assembly_name:Contig19"
27751. 27850
/estimated_length=unknown
27851. 29522
/note="assembly_name:Contig20"
29523. 29622
vector_side:left"
clone_end:SP6
misc_feature
29623. 33750
```

```
gap
/note="assembly_name:Contig21"
33751. 33850
/estimated_length=unknown
misc_feature
33851. 37904
/note="assembly_name:Contig22"
clone_end:r7
vector_side:right"
gap
37905. 38004
/estimated_length=unknown
misc_feature
38005. 43341
/note="assembly_name:Contig23"
43342. 43341
/estimated_length=unknown
43342. 50634
/note="assembly_name:Contig24"
50635. 50734
/estimated_length=unknown
50735. 59411
/note="assembly_name:Contig25"
59412. 59511
/estimated_length=unknown
59512. 68503
/note="assembly_name:Contig26"
68504. 68603
/estimated_length=unknown
68604. 79177
/note="assembly_name:Contig27"
79178. 79277
/estimated_length=unknown
79278. 86283
/note="assembly_name:Contig28"
86284. 86383
/estimated_length=unknown
86384. 96915
/note="assembly_name:Contig29"
96916. 97015
/estimated_length=unknown
97016. 107357
/note="assembly_name:Contig30"
107358. 107457
/estimated_length=unknown
107458. 119546
/note="assembly_name:Contig31"
119547. 119646
/estimated_length=unknown
119647. 134223
/note="assembly_name:Contig32"
134224. 134323
/estimated_length=unknown
134324. 147018
misc_feature

Query Match 1.7%; Score 53; DB 14; Length 179462;
Best Local Similarity 100.0%; Pred. No. 1.6e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2888 TGAGGACAGTGGATACCTGAGGACGAGGATTGAGACGAGCTGGCCACAT 2940
DB 74903 TGAGGACAGTGGATACCTGAGGACGAGGATTGAGACGAGCTGGCCACAT 74955

RESULT 32 AC007332 185000 bp DNA linear HTG 14-DEC-1999
AC007332 Homo sapiens chromosome 4, *** SEQUENCING IN PROGRESS ***, 14
LOCUS AC007332
DEFINITION HTG; HTGS PHASE1.
ACCESSION AC007332.11 GI:573831
VERSION HTG; HTGS PHASE1.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
```

REFERENCE 1 (bases 1 to 185000)  
 AUTHORS Stone,N.E.,Schmutz,J.J., Cox,D.R. and Myers,R.M.  
 TITLE Direct Submission  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 185000)  
 AUTHORS Stone,N.E.,Schmutz,J.J., Cox,D.R. and Myers,R.M.  
 TITLE Direct Submission  
 JOURNAL Submitted (20-APR-1999) Department of Genetics, Stanford Human Genome Center, 865 California Avenue, Palo Alto, CA 94304, USA  
 On Dec 14, 1999 this sequence version replaced gi:6137906.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 14 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

1 1233: contig of 1233 bp in length  
 \* 1234 1994: gap of unknown length  
 \* 1395 3971: contig of 1977 bp in length  
 \* 3972 4732: gap of unknown length  
 \* 4733 6959: contig of 2227 bp in length  
 \* 6960 7720: gap of unknown length  
 \* 7721 9753: contig of 2033 bp in length  
 \* 9754 10514: gap of unknown length  
 \* 10515 12313: contig of 1799 bp in length  
 \* 12314 13074: gap of unknown length  
 \* 13075 15817: contig of 2743 bp in length  
 \* 15818 16578: gap of unknown length  
 \* 16579 21924: contig of 5346 bp in length  
 \* 21925 22685: gap of unknown length  
 \* 22686 27521: contig of 4836 bp in length  
 \* 27522 28282: gap of unknown length  
 \* 28283 33545: contig of 5263 bp in length  
 \* 33546 51372: contig of 17066 bp in length  
 \* 51373 52133: gap of unknown length  
 \* 52134 73810: contig of 21677 bp in length  
 \* 73811 74570: gap of unknown length  
 \* 74571 101382: contig of 26812 bp in length  
 \* 101383 102142: gap of unknown length  
 \* 102143 143256: contig of 41114 bp in length  
 \* 143257 144016: gap of unknown length  
 \* 144017 185000: contig of 40984 bp in length.

FEATURES  
 source  
 1. .185000  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="4"  
 1234. .1994  
 /estimated\_length=unknown  
 3972. .4732  
 /estimated\_length=unknown  
 6960. .7720  
 /estimated\_length=unknown  
 9754. .10514  
 /estimated\_length=unknown  
 12314. .13074  
 /estimated\_length=unknown  
 15818. .16578  
 /estimated\_length=unknown  
 21925. .22685  
 /estimated\_length=unknown  
 27522. .28282  
 /estimated\_length=unknown  
 33546. .34306  
 /estimated\_length=unknown  
 51373. .52133  
 /estimated\_length=unknown  
 73811. .74570  
 /estimated\_length=unknown

gap 101383. .102142  
 /estimated\_length=unknown  
 gap 143257. .144016  
 /estimated\_length=unknown

ORIGIN  
 Query Match 1.7%; Score 53; DB 14; Length 185000;  
 Best Local Similarity 100.0%; Pred.No. 1.6e-16;  
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2888 TTAGGACAGGTGATCACTGAGCCAGAGATTGAGACCAAGCTGGCCCAACT 2940  
 Db 168527 TTAGGACAGGTGATCACTGAGCCAGAGATTGAGACCAAGCTGGCCCAACT 168579

RESULT 33  
 AP001333  
 LOCUS  
 DEFINITION Homo sapiens chromosome 18 clone RP11-675J18 map 18q12, WORKING  
 DRAFT SEQUENCE, 25 unordered pieces.  
 ACCESSION AP001333 GI:9186910  
 VERSION AP001333.3  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 1 (bases 1 to 189181)  
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,  
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.  
 Homo sapiens 189,181 genomic DNA of 18q12  
 Published Only in Database (2000)  
 2 (bases 1 to 189181)  
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,  
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.  
 Direct Submission  
 Submitted (02-MAR-2000) Maashira Hattori, The Institute of Physical  
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),  
 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555, Japan  
 (E-mail:hattori@gs.c.riken.go.jp, URL:http://hsp.gsc.riken.go.jp/  
 Tel:81-42-778-9923, Fax:81-42-778-9924)  
 On Jul 14, 2000 this sequence version replaced gi:8117256.

COMMENT  
 ----- Genome Center  
 Center: RIKEN Genomic Sciences Center(GSC)  
 Center code: RIKEN  
 Web site: http://hsp.gsc.riken.go.jp/  
 Contact: hattori@gs.c.riken.go.jp  
 ----- Project Information  
 Center project name: HumDraff18  
 Center clone name: RP11-675J18  
 ----- Summary Statistics  
 Sequencing vector: PCR products; 100% of reads  
 Chemistry: Dye-terminator ET-amersham; 100% of reads Assembly  
 Program: Phrap; version 0.990329  
 Consensus quality: 181064 bases at least Q40  
 Consensus quality: 183725 bases at least Q30  
 Consensus quality: 185304 bases at least Q20  
 Insert size: 186781; sum-of-contigs  
 Quality coverage: 9.17x in Q20 bases; sum-of-contigs  
 -----  
 NOTE: This is a 'working draft' sequence. It currently consists of  
 25 contigs. The true order of the pieces is not known and their  
 order in this sequence record is arbitrary. Gaps between the  
 contigs are represented as runs N, but the exact sizes of the gaps  
 are unknown. This record will be updated with the finished sequence  
 as soon as it is available and the accession number will be  
 preserved

1 37335 contig of 37335 bp in length  
 37436 67386 contig of 29951 bp in length  
 67467 90245 contig of 22759 bp in length  
 90346 100199 contig of 9854 bp in length  
 100300 108800 contig of 8501 bp in length

```

108901 117127 contig of 8227 bp in length
117228 124504 contig of 7277 bp in length
124605 132198 contig of 7594 bp in length
132299 138051 contig of 5753 bp in length
138152 144498 contig of 6347 bp in length
144599 150308 contig of 5710 bp in length
150409 154983 contig of 4575 bp in length
155084 158506 contig of 3423 bp in length
158607 162373 contig of 3767 bp in length
162474 164257 contig of 1784 bp in length
164358 167268 contig of 2911 bp in length
167369 171012 contig of 3644 bp in length
171113 174768 contig of 3656 bp in length
174869 177195 contig of 2327 bp in length
177296 179508 contig of 2213 bp in length
179609 181405 contig of 1797 bp in length
181506 184223 contig of 2718 bp in length
184324 185884 contig of 1561 bp in length
185985 187859 contig of 1875 bp in length
187960 189181 contig of 1222 bp in length.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 37335: contig of 37335 bp in length
* 37336 37435: gap of 100 bp
* 37436 67386: contig of 29951 bp in length
* 67387 67487: gap of 100 bp
* 67487 90245: contig of 22759 bp in length
* 90246 90345: gap of 100 bp
* 90345 100199: contig of 9854 bp in length
* 100200 100299: gap of 100 bp
* 100300 108800: contig of 8501 bp in length
* 108801 108900: gap of 100 bp
* 108901 117127: contig of 8227 bp in length
* 117128 117227: gap of 100 bp
* 117228 124504: contig of 7277 bp in length
* 124505 124604: gap of 100 bp
* 124605 132198: contig of 7534 bp in length
* 132199 132298: gap of 100 bp
* 132299 138051: contig of 5753 bp in length
* 138052 138151: gap of 100 bp
* 138152 144498: contig of 6347 bp in length
* 144499 144598: gap of 100 bp
* 144599 150308: contig of 5710 bp in length
* 150309 150408: gap of 100 bp
* 150409 154983: contig of 4575 bp in length
* 154984 155083: gap of 100 bp
* 155084 158506: contig of 3423 bp in length
* 158507 158606: gap of 100 bp
* 158607 162373: contig of 3767 bp in length
* 162374 162473: gap of 100 bp
* 162474 164257: contig of 1784 bp in length
* 164258 164357: gap of 100 bp
* 164358 167268: contig of 2911 bp in length
* 167269 167368: gap of 100 bp
* 167369 171012: contig of 3644 bp in length
* 171113 171112: gap of 100 bp
* 171113 174768: contig of 3656 bp in length
* 174769 174868: gap of 100 bp
* 174869 177195: contig of 2327 bp in length
* 177296 179508: gap of 100 bp
* 179509 179608: contig of 2213 bp in length
* 179609 181405: gap of 100 bp
* 181406 181505: gap of 100 bp
* 181506 184223: contig of 2718 bp in length
* 184224 184323: gap of 100 bp
* 184324 185884: contig of 1561 bp in length

```

```

* 185885 185984: gap of 100 bp
* 185985 187859: contig of 1875 bp in length
* 187860 187959: gap of 100 bp
* 187960 189181: contig of 1222 bp in length.
location/Qualifiers
1. 189181
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="18"
/map="18q12"
/clone="RP11-675J18"
1. 37335

```

```

misc_feature /note="assembly_fragment" 37336..67386
misc_feature /note="assembly_fragment" 67487..90245
misc_feature /note="assembly_fragment" 90346..100199
misc_feature /note="assembly_fragment_clone_end:T7_vector_side:left" 100300..108800
misc_feature /note="assembly_fragment" 108901..117127
misc_feature /note="assembly_fragment" 117228..124504
misc_feature /note="assembly_fragment" 124605..132198
misc_feature /note="assembly_fragment" 132299..138051
misc_feature /note="assembly_fragment" 138152..144498
misc_feature /note="assembly_fragment" 144599..150308
misc_feature /note="assembly_fragment" 150409..154983
misc_feature /note="assembly_fragment" 155084..158506
misc_feature /note="assembly_fragment" 158607..162373
misc_feature /note="assembly_fragment" 162474..164257
misc_feature /note="assembly_fragment_clone_end:SP6_vector_side:left" 164358..167268
misc_feature /note="assembly_fragment" 167369..171012
misc_feature /note="assembly_fragment" 171113..174768
misc_feature /note="assembly_fragment" 174869..177195
misc_feature /note="assembly_fragment" 177296..179508
misc_feature /note="assembly_fragment" 179609..181405
misc_feature /note="assembly_fragment" 181506..184223
misc_feature /note="assembly_fragment" 184324..185884
misc_feature /note="assembly_fragment" 185985..187859
misc_feature /note="assembly_fragment" 187960..189181
misc_feature /note="assembly_fragment"

```

## ORIGIN

```

Query Match 1.7%; Score 53; DB 14; Length 189181;
Best Local Similarity 100.0%; Pred. No. 1.6e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2888 TTAGGCAAGTGTATCACTTGAAGCCAGAGACTTCGAGACGCTGCGCAACT 2940
DB 137657 TGAAGCAGGTGATCACTTGAAGCCAGAGACTTCGAGACGCTGCGCAACT 137709

```



## COMMENT

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Mar 9, 2002 this sequence version replaced gi:15799649.

----- Genome Center

Center: Washington University Genome Sequencing Center

Center code: WUGSC

Web site: <http://genome.wustl.edu/gsc>

Contact: [saplens@wustl.wustl.edu](mailto:saplens@wustl.wustl.edu)

----- Summary Statistics

Center project name: H\_NH0326119

-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

## MAPPING INFORMATION:

Mapping information for this clone was provided by Dr. John D. McPherson, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu/gsc>

## SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osoegawa, K., Moon, P.Y., Zhao, B., Frengen, B., Tatemio, M., Catelese, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBACe3.6

## NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP11-274B16; the clone sequenced to the right is RP11-386I19. Actual start of this clone is at base position 1 of RP11-326I19; actual end is at base position 203583 of RP11-326I19.

There are single plasmid subclone regions from 9433 to 9523, 179086 to 179114 and 195823 to 195835. There is an unresolved homopolymeric run from 74097 to 74101. Polymorphisms have been identified between AC025084, AC108139, AC007332 and AC092437. Data from AC025084, AC108139 and AC007332 was used to finish this clone, AC092437.

## FEATURES

## source

Location/Qualifiers  
1..203583  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="4"  
/map="4"  
/clone="RP11-326I19"  
/clone\_1fb="RPCI-11"  
382..508  
/rpt\_family="L2"  
repeat\_region  
522..639  
/rpt\_family="MIR"  
repeat\_region  
1049..1187  
/rpt\_family="MIR"  
repeat\_region  
1237..1347  
/rpt\_family="L2"  
repeat\_region  
1350..1389  
/rpt\_family="L2"  
repeat\_region  
1485..1559  
/rpt\_family="MIR"

repeat\_region 1640..2055  
/rpt\_family="MALR"  
repeat\_region 2254..2391  
/rpt\_family="MIR"  
repeat\_region 2717..3003  
/rpt\_family="L1"  
repeat\_region 3047..3709  
/rpt\_family="MER2\_type"  
repeat\_region 3732..3781  
/rpt\_family="CATATA)n"  
repeat\_region 3802..3929  
/rpt\_family="L1"  
repeat\_region 3935..3961  
/rpt\_family="AT-rich"  
repeat\_region 4389..4431  
/rpt\_family="(CA)n"  
repeat\_region 5193..5266  
/rpt\_family="L2"  
repeat\_region 5556..5645  
/rpt\_family="MIR"  
repeat\_region 5999..6169  
/rpt\_family="MER1\_type"  
repeat\_region 6202..6316  
/rpt\_family="MIR"  
repeat\_region 6555..6606  
/rpt\_family="MIR"  
repeat\_region 6910..7272  
/rpt\_family="MALR"  
repeat\_region 7537..7905  
/rpt\_family="MALR"  
repeat\_region 7909..9510  
/rpt\_family="MALR"  
repeat\_region 9523..9900  
/rpt\_family="MALR"  
repeat\_region 9925..10132  
/rpt\_family="MALR"  
repeat\_region 10133..10268  
/rpt\_family="Alu"  
repeat\_region 10269..10411  
/rpt\_family="MALR"  
repeat\_region 11024..11061  
/rpt\_family="L2"  
repeat\_region 11123..11200  
/rpt\_family="MIR"  
repeat\_region 11236..11520  
/rpt\_family="L2"  
repeat\_region 13157..13347  
/rpt\_family="MIR"  
repeat\_region 13360..13772  
/rpt\_family="ATGGTG)n"  
repeat\_region 13754..13929  
/rpt\_family="(TGG)n"  
repeat\_region 13931..13977  
/rpt\_family="(ATG)n"  
repeat\_region 13987..14322  
/rpt\_family="MIR"  
repeat\_region 14423..14527  
/rpt\_family="L2"  
repeat\_region 14594..14708  
/rpt\_family="L2"  
repeat\_region 14855..15035  
/rpt\_family="MIR"  
repeat\_region 15426..15727  
/rpt\_family="Alu"  
repeat\_region 16021..16126  
/rpt\_family="MIR"  
repeat\_region 16429..16856  
/rpt\_family="MALR"  
repeat\_region 17105..17215  
/rpt\_family="MIR"  
repeat\_region 17225..17460  
/rpt\_family="L1"  
repeat\_region 17461..17820

```

repeat_region      /rpt_family="MaLR"
                    17821..19578
repeat_region      /rpt_family="L1"
                    19582..20133
repeat_region      /rpt_family="L1"
                    20134..20429
repeat_region      /rpt_family="Alu"
                    20430..21136
repeat_region      /rpt_family="L1"
                    21282..21533
repeat_region      /rpt_family="MIR"
                    22088..22402
repeat_region      /rpt_family="Alu"
                    22757..23239
repeat_region      /rpt_family="MaLR"
                    23406..23686
repeat_region      /rpt_family="ERV1"
                    24004..24189
repeat_region      /rpt_family="MaLR"
                    24320..24487
repeat_region      /rpt_family="MIR"
                    24514..24601
repeat_region      /rpt_family="L2"
                    25075..25182
repeat_region      /rpt_family="MIR"

```

```

Query Match      1.7%; Score 53; DB 8; Length 203583;
Best Local Similarity 100.0%; Pred. No. 1.6e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

OY      2888 TGAGGAGGTGATGACCTGAGGCGAGGATTGAGACGAGCTGCGCAACAT 2940
Db      130670 TGAAGCAGTGTGATCTCCTGAGGCGAGGATTGAGACGAGCTGCGCAACAT 130722

```

```

RESULT 36
CR847805      204373 bp      DNA      linear      HTG 05-NOV-2004
LOCUS      CR847805      Homo sapiens chromosome 6 clone DMM-118E22, 24 unordered pieces.
DEFINITION
ACCESSION
VERSION      CR847805.4      GI:55468491
KEYWORDS      HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
Sims.S.
1 (bases 1 to 204373)
Direct Submission
Submitted (04-NOV-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Nov 5, 2004 this sequence version replaced gi:55294819.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BM118E22
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 10% of reads
Consensus quality: 195146 bases at least Q40
Consensus quality: 198151 bases at least Q30
Consensus quality: 199848 bases at least Q20
Insert size: 202073; sum-of-contigs
Insert size: 99271; 1.0% error; agarose-ep
Quality coverage: 6.13x in Q20 bases; sum-of-contigs Quality
coverage: 14.67x in Q20 bases; agarose-ep

```

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 24 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

1 3167: contig of 3167 bp in length
3168 3267: gap of 100 bp
3268 5570: contig of 2303 bp in length
5571 5670: gap of 100 bp
5671 13473: contig of 7803 bp in length
13474 13574: gap of 100 bp
13574 17619: contig of 4046 bp in length
17620 17719: gap of 100 bp
17720 20589: contig of 2870 bp in length
20590 20689: gap of 100 bp
20690 35318: contig of 14629 bp in length
35319 35418: gap of 100 bp
35419 40521: contig of 5103 bp in length
40522 40621: gap of 100 bp
40622 53837: contig of 1316 bp in length
53838 53937: gap of 100 bp
53938 56328: contig of 2391 bp in length
56329 56428: gap of 100 bp
56429 60179: contig of 3751 bp in length
60180 60279: gap of 100 bp
60280 63603: contig of 3324 bp in length
63604 63703: gap of 100 bp
63704 69568: contig of 5865 bp in length
69569 72492: contig of 2824 bp in length
72493 72592: gap of 100 bp
72593 74831: contig of 2239 bp in length
74832 74931: gap of 100 bp
74932 77561: contig of 2630 bp in length
77562 81419: gap of 100 bp
81419 81519: contig of 378 bp in length
81520 86554: contig of 515 bp in length
86555 86754: gap of 100 bp
86755 89390: contig of 2636 bp in length
89391 89490: gap of 100 bp
89491 94282: contig of 4792 bp in length
94283 94382: gap of 100 bp
94383 102792: contig of 8410 bp in length
102793 112442: gap of 100 bp
112443 112442: contig of 9550 bp in length
112443 112542: gap of 100 bp
112543 132047: contig of 19505 bp in length
132048 132147: gap of 100 bp
132148 162842: contig of 30695 bp in length
162843 204373: contig of 41431 bp in length.
162943

```

```

FEATURES
source      1..204373
             /organism="Homo sapiens"
             /mol_type="Genomic DNA"
             /db_xref="taxon:9606"
             /chromosome="6"
             /clone.lib="DMM-118E22"

```

```

misc_feature      1..3167
                  /note="assembly_fragment:00230
                  fragment_chain:1"

```

```

misc_feature      3268..5570
                  /note="assembly_fragment:00124
                  fragment_chain:1"

```

```

misc_feature      5671..13473
                  /note="assembly_fragment:00481
                  fragment_chain:1"

```

```

misc_feature      13574..17619
                  /note="assembly_fragment:00293

```



```

misc_feature      fragment chain:1"
                  17720..20589
                  /note="assembly_fragment:00278
                  fragment_chain:2"
misc_feature      20690..35318
                  /note="assembly_fragment:00574
                  fragment_chain:2"
misc_feature      35419..40521
                  /note="assembly_fragment:00528
                  fragment_chain:2"
misc_feature      40622..53837
                  /note="assembly_fragment:00692
                  fragment_chain:2"
misc_feature      53938..56328
                  /note="assembly_fragment:00211
                  fragment_chain:3"
misc_feature      56429..60179
                  /note="assembly_fragment:00341
                  fragment_chain:3"
misc_feature      60280..63603
                  /note="assembly_fragment:00252
                  fragment_chain:4"
misc_feature      63704..69568
                  /note="assembly_fragment:00406
                  fragment_chain:4"
misc_feature      69669..72492
                  /note="assembly_fragment:00220"
                  fragment_chain:4"
misc_feature      72593..74831
                  /note="assembly_fragment:00241"
                  fragment_chain:4"
misc_feature      74932..77561
                  /note="assembly_fragment:00265"
                  fragment_chain:4"
misc_feature      77662..81419
                  /note="assembly_fragment:00308"
                  fragment_chain:4"
misc_feature      81520..86654
                  /note="assembly_fragment:00324"
                  fragment_chain:4"
misc_feature      86755..89390
                  /note="assembly_fragment:00358"
                  fragment_chain:4"
misc_feature      89491..94282
                  /note="assembly_fragment:0036"
                  fragment_chain:4"
misc_feature      94383..102792
                  /note="assembly_fragment:00437"
                  fragment_chain:4"
misc_feature      102893..112442
                  /note="assembly_fragment:00840.0"
                  fragment_chain:4"
misc_feature      112543..133047
                  /note="assembly_fragment:01041.0"
                  fragment_chain:4"
misc_feature      132148..162842
                  /note="assembly_fragment:01254.0"
                  fragment_chain:4"
misc_feature      162943..204373
                  /note="assembly_fragment:01615.0"
                  fragment_chain:4"
ORIGIN
Query Match      1.7%; Score 53; DB 14; Length 204373;
Best Local Similarity 100.0%; Pred. No. 1.6e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      2888 TGAGCAGGTGATCCTGAGGCCAGAGTTTCAGACAGCTGGCCAACT 2940
DB      156538 TGAGCAGGTGATCCTGAGGCCAGAGTTTCAGACAGCTGGCCAACT 156590

```

```

AUTHORS      DOE Joint Genome Institute, Stanford Human Genome Center and Los
              Alamos National Laboratory.
TITLE      Direct Submission
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 208220)
AUTHORS      DOE Joint Genome Institute.
TITLE      Direct Submission
JOURNAL      Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
              Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE    3 (bases 1 to 208220)
AUTHORS      DOE Joint Genome Institute.
TITLE      Direct Submission
JOURNAL      Submitted (04-DEC-2001) Production Sequencing Facility, DOE Joint
              Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE    4 (bases 1 to 208220)
AUTHORS      DOE Joint Genome Institute, Stanford Human Genome Center and Los
              Alamos National Laboratory.
TITLE      Direct Submission
JOURNAL      Submitted (14-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell
              Drive, Walnut Creek, CA 94598, USA
COMMENT      On Mar 14, 2003 this sequence version replaced gi:17298588.
              Draft Sequence Produced by DOE Joint Genome Institute
              www.jgi.doe.gov
              National Laboratory
              www.sbgc.stanford.edu
              Quality: Phrap Quality >=40 100% of Sequence;
              Estimated Total Number of Errors is 0.
FEATURES
source
1..208220
location/Qualifiers
    /organism="Homo sapiens"
    /mol_type="Genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="16"
    /clone="RP11-459F6"
ORIGIN
Query Match      1.7%; Score 53; DB 8; Length 208220;
Best Local Similarity 100.0%; Pred. No. 1.6e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      2888 TGAGCAGGTGATCCTGAGGCCAGAGTTTCAGACAGCTGGCCAACT 2940
DB      14755 TGAGCAGGTGATCCTGAGGCCAGAGTTTCAGACAGCTGGCCAACT 14807

```

```

RESULT 38
LOCUS      AC092434      218872 bp      DNA      linear      HTG 27-JUL-2001
DEFINITION Homo sapiens chromosome 4 clone RP11-210L21, *** SEQUENCING IN
PROGRESS ***, 40 unordered pieces.
ACCESSION  AC092434
VERSION     AC092434.1 GI:14595956
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.
REFERENCE   1 (bases 1 to 218872)
AUTHORS    Waterston,R.H.
TITLE      The sequence of Homo sapiens clone
JOURNAL    Unpublished
REFERENCE   2 (bases 1 to 218872)
AUTHORS    Waterston,R.H.
TITLE      Direct Submission
JOURNAL    Submitted (04-JUL-2001) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC

```





CR753824  
 LOCUS CR753824 323185 bp DNA linear HTG 05-NOV-2004  
 DEFINITION Homo sapiens chromosome 6 clone DADB-56024, \*\*\* SEQUENCING IN PROGRESS \*\*\*; 4 unoriented pieces.  
 ACCESSION CR753824  
 VERSION CR753824.2 GI:55468370  
 KEYWORDS HTG, HTGS PHASE1, HTGS\_ACTIVEFIN, HTGS\_FULLTOP.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 323185)  
 Leongamornlert,D.  
 Direct Submission  
 Submitted (04-NOV-2004) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk  
 Clone requests: clonerequests@sanger.ac.uk  
 On Nov 5, 2004 this sequence version replaced gi:51571640.  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: humquery@sanger.ac.uk  
 ----- Project Information  
 Center project name: hdb56024  
 ----- Summary Statistics  
 Assembly program: XGAP4; version 4.5  
 Chemistry: Dye-terminator; 100% of reads  
 Consensus quality: 89730 bases at least Q40  
 Consensus quality: 89774 bases at least Q30  
 Consensus quality: 89790 bases at least Q20  
 Insert size: 322885; sum-of-contigs  
 Insert size: 96449; 1.4% error; agarose-fp  
 Quality coverage: 1.82x in Q20 bases; sum-of-contigs quality coverage: 6.22x in Q20 bases; agarose-fp  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.  
 \* 1 15449: contig of 15449 bp in length  
 \* 15450 15549: gap of 100 bp  
 \* 15550 60747: contig of 45198 bp in length  
 \* 60748 60847: gap of 100 bp  
 \* 60848 165052: contig of 104205 bp in length  
 \* 165053 165152: gap of 100 bp  
 \* 165153 323185: contig of 158033 bp in length.  
 Location/Qualifiers  
 1..323185  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="6"  
 /clone="DADB-56024"  
 /clone\_1fb="DNA-arcs-BAC.1-DBB.1"  
 1..15449  
 /note="assembly fragment:00142  
 fragment\_chain:1"  
 15550..60747  
 /note="assembly fragment:00419  
 fragment\_chain:1"  
 60848..165052  
 /note="assembly fragment:00920"  
 165153..323185  
 /note="assembly fragment:00949"  
 ORIGIN  
 Query Match 1.7%; Score 53; DB 14; Length 323185;

Best Local Similarity 100.0%; Pred. No. 1.6e-16;  
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2888 TGAGGACAGTGGATCACTGAGGCCAGAGAGTTCCAGACAGCTGCGCAACT 2940  
 Db 288592 TGAGGACAGTGGATCACTGAGGCCAGAGAGTTCCAGACAGCTGCGCAACT 288644

RESULT 41  
 LOCUS AP001710 340000 bp DNA linear PRI 21-MAY-2003  
 DEFINITION Homo sapiens genomic DNA, chromosome 21q, section 54/105.  
 ACCESSION AP001710 AL163255 BA000005  
 VERSION AP001710.1 GI:7768781  
 KEYWORDS  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
 1  
 Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T., Park,H.S., Toyoda,A., Ishii,K., Tokoi,Y., Choi,D.K., Soeda,E., Ohki,M., Takagi,T., Sakaki,Y., Tauden,S., Blechschmidt,K., Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R., Patterson,D., Reichwald,K., Rump,A., Schillabel,M., Schudy,A., Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K., Asakawa,S., Shintani,A., Sasaki,K., Nagamine,K., Mitsuyama,S., Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordstiek,G., Hornischner,K., Barandt,P., Scharfe,M., Schoen,O., Desario,A., Reichelt,J., Kauer,G., Bloeker,H., Ramser,J., Beck,A., Klages,S., Hennig,S., Riesemann,L., Daggand,E., Wehrmeyer,S., Borzym,K., Gardiner,K., Nizetic,D., Francis,F., Lehnach,H., Reinhardt,R. and Yaspo,M.L.  
 The DNA sequence of human chromosome 21  
 Nature 405 (6784), 311-319 (2000)  
 10830953  
 2 (bases 1 to 340000)  
 Hattori,M., Fujiyama,A., Taylor,T.D., Watanabe,H., Yada,T., Park,H.S., Toyoda,A., Ishii,K., Tokoi,Y., Choi,D.K., Soeda,E., Ohki,M., Takagi,T., Sakaki,Y., Tauden,S., Blechschmidt,K., Polley,A., Menzel,U., Delabar,J., Kumpf,K., Lehmann,R., Patterson,D., Reichwald,K., Rump,A., Schillabel,M., Schudy,A., Zimmermann,W., Rosenthal,A., Kudoh,J., Shibuya,K., Kawasaki,K., Asakawa,S., Shintani,A., Sasaki,T., Nagamine,K., Mitsuyama,S., Antonarakis,S.E., Minoshima,S., Shimizu,N., Nordstiek,G., Hornischner,K., Barandt,P., Scharfe,M., Schoen,O., Desario,A., Reichelt,J., Kauer,G., Bloeker,H., Ramser,J., Beck,A., Klages,S., Hennig,S., Riesemann,L., Daggand,E., Wehrmeyer,S., Borzym,K., Gardiner,K., Nizetic,D., Francis,F., Lehnach,H., Reinhardt,R. and Yaspo,M.L.  
 Direct Submission  
 Submitted (10-APR-2000) The Chromosome 21 Mapping and Sequencing Consortium \* RIKEN Genomic Sciences Center, Human Genome Research Group \* Institute of Molecular Biotechnology, Genome Analysis \* Keio University School of Medicine, Dept. of Molecular Biology \* GBF, Dept. of Genome Analysis \* Max-Planck Institute for Molecular Genetics (addresses see below)  
 On May 30, 2000 this sequence version replaced gi:7717322.  
 The chromosome 21 mapping and sequencing consortium consisting of \* RIKEN Genomic Sciences Center, Human Genome Research Group, \* Sagamihara 228-8555, Japan.  
 \* e-mail: hattori@gsc.riken.go.jp  
 \* URL: http://hsp.gsc.riken.go.jp/  
 and  
 \* Institute of Molecular Biotechnology, Genome Analysis, \* Beutenbergstrasse 11, D-07745 Jena, Germany,  
 \* e-mail: gscj-submit@genome.imb-jena.de  
 \* URL: http://genome.imb-jena.de/  
 and  
 \* Keio University School of Medicine, Molecular Biology, \* Tokyo 160-8582, Japan,  
 \* e-mail: nshimizu@mb-med.keio.ac.jp

\* URL: <http://www.dmb.med.keio.ac.jp/>  
and  
\* GBF, Dept. of Genome Analysis,  
\* Mascheroder Weg 1, D-38124 Braunschweig, Germany, \* e.mail:  
\* info.genome@gbf.de  
\* URL: <http://genome.gbf.de/>  
and  
\* Max-Planck Institute for Molecular Genetics,  
\* Innesstrasse 73, D-14195 Berlin, Germany,  
\* e.mail: [info-chr21emoigen.mpg.de](mailto:info-chr21emoigen.mpg.de)  
\* URL: <http://chr21.rz-berlin.mpg.de/>  
AL163255: Submitted (10-Apr-2000).  
Location/Qualifiers

## FEATURES

source 1..340000  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="21"  
1..52327  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.1"  
/clone="R14R21, 5' partial"  
/clone\_1lb="RPC1-11 BAC library"  
/note="Accession No. AP000245"  
51884..146095  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.1"  
/clone="R137B7"  
/clone\_1lb="RPC1-11 BAC library"  
/note="Accession No. AP000246"  
145455..259373  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.1"  
/clone="PAC871G18"  
/clone\_1lb="RPS PAC library"  
/note="Accession No. AP000247"  
256917..>340000  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="21"  
/map="21q22.1"  
/clone="R289H18, 3' partial"  
/clone\_1lb="RPC1-11 BAC library"  
/note="Accession No. AP000248"  
complement(783..1089)  
/note="AluSc"  
/rpt\_family="SINE/Alu"  
/rpt\_type=DISPERSED  
complement(3359..3661)  
/note="AluUo"  
/rpt\_family="SINE/Alu"  
/rpt\_type=DISPERSED  
4287..4321  
/note="(?)n"  
/rpt\_family="Simple\_repeat"  
/rpt\_type=TANDEM  
5078..5654  
/note="MER74A"  
/rpt\_family="LTR/MER73-group"  
/rpt\_type=DISPERSED  
6855..7314  
/note="LTR2"  
/rpt\_family="LTR/Retroviral"

repeat\_region /rpt\_type=DISPERSED  
7666..7725  
/note="(TA)n"  
/rpt\_family="Simple\_repeat"  
repeat\_region /rpt\_type=TANDEM  
7761..7825  
/note="(TA)n"  
/rpt\_family="Simple\_repeat"  
repeat\_region /rpt\_type=TANDEM  
8223..8411  
/note="L1M4"  
/rpt\_family="LINE/L1"  
repeat\_region /rpt\_type=DISPERSED  
8625..9032  
/note="L1M4"  
/rpt\_family="LINE/L1"  
repeat\_region /rpt\_type=DISPERSED  
9504..10065  
/note="L1ME3"  
/rpt\_family="LINE/L1"  
repeat\_region /rpt\_type=DISPERSED  
complement(10066..10198)  
/note="FLAM C"  
/rpt\_family="SINE/Alu"  
repeat\_region /rpt\_type=DISPERSED  
10199..10664  
/note="L1ME3"  
/rpt\_family="LINE/L1"  
repeat\_region /rpt\_type=DISPERSED  
10690..10950  
/note="L1ME3"  
/rpt\_family="LINE/L1"  
repeat\_region /rpt\_type=DISPERSED  
11012..11310  
/note="AluY"  
/rpt\_family="SINE/Alu"  
repeat\_region /rpt\_type=DISPERSED  
11596..11903  
/note="AluY"  
/rpt\_family="SINE/Alu"  
repeat\_region /rpt\_type=DISPERSED  
complement(11311..11475)  
/note="AluY"  
/rpt\_family="SINE/Alu"  
repeat\_region /rpt\_type=DISPERSED  
complement(11377..14219)  
/note="MLT1B2"  
/rpt\_family="LTR/MaLR"  
repeat\_region /rpt\_type=DISPERSED  
14220..14505  
/note="AluY"  
/rpt\_family="SINE/Alu"  
repeat\_region /rpt\_type=DISPERSED  
complement(14506..14554)  
/note="MLT1B2"  
/rpt\_family="LTR/MaLR"  
repeat\_region /rpt\_type=DISPERSED  
14670..14795  
/note="MIR"  
/rpt\_family="SINE/MIR"  
repeat\_region /rpt\_type=DISPERSED  
14838..14909  
/note="L2"  
/rpt\_family="LINE/L2"

Query Match 1.7%; Score 53; DB 8; length 340000;  
Best local similarity 100.0%; Pred.No.1.6e-16;  
Matches 53; conservative 0; mismatches 0; indels 0; gaps 0;

Qy 2888 TGAGCAGGTGATCACTGAGCCGAGAGTTGAGACCAAGCTTGGCCAACT 2940  
DB 159396 TGAGCAGGTGATCACTGAGCCGAGAGTTGAGACCAAGCTTGGCCAACT 159448

RESULT 42  
LOCUS CS039425 349980 bp DNA linear PAT 22-MAR-2005  
DEFINITION Sequence 235 from Patent WO2005019477.  
ACCESSION CS039425  
VERSION CS039425.1 GI:61847452  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE  
1 Lewin,J., Berlin,K., Hildmann,T., Olek,A., Beck,S. and Novik,K.  
AUTHORS Methods and compositions for differentiating classes of cell types  
TITLE using epigenetic markers  
JOURNAL Patent: WO 2005019477-A 235 03-MAR-2005;  
EpiGenomics AG (DE)  
FEATURES  
source  
1.349980 Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
/note="unknown base"  
seq 205 to long, 4.267.840  
replaced by  
new seq 205, from 0.000.001 to 0.349.980  
new seq 222, from 0.300.001 to 0.649.980  
new seq 223, from 0.600.001 to 0.949.980  
new seq 224, from 0.900.001 to 1.249.980  
new seq 225, from 1.200.001 to 1.549.980  
new seq 226, from 1.500.001 to 1.849.980  
new seq 227, from 1.800.001 to 2.149.980  
new seq 228, from 2.100.001 to 2.449.980  
new seq 229, from 2.400.001 to 2.749.980  
new seq 230, from 2.700.001 to 3.049.980  
new seq 231, from 3.000.001 to 3.349.980  
new seq 232, from 3.300.001 to 3.649.980  
new seq 233, from 3.600.001 to 3.949.980  
new seq 234, from 3.900.001 to 4.249.980  
new seq 235, from 4.200.001 to 4.549.980  
new seq 236, from 4.500.001 to 4.647.455"  
ORIGIN  
Query Match 1.7%; Score 53; DB 6; Length 349980;  
Best Local Similarity 100.0%; Pred. No. 1.6e-16;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2888 TGAGGCGATGATCCTGAGGCGAGGCTCGAGACCAAGCTGGCCAACT 2940  
DB 240219 TGAGGCGATGATCCTGAGGCGAGGCTCGAGACCAAGCTGGCCAACT 240271  
RESULT 43  
LOCUS BD027540/c 381 bp DNA linear PAT 27-AUG-2002  
DEFINITION Sequence tag and encoded human protein.  
ACCESSION BD027540  
VERSION BD027540.1 GI:22569282  
KEYWORDS JP 2001269182-A/3786.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 381)  
AUTHORS Edwards,J.B.D.M., Ducclair,E. and Jordan,J.Y.  
TITLE Sequence tag and encoded human protein  
JOURNAL Patent: JP 2001269182-A 3786 02-OCT-2001;  
GENSET  
COMMENT OS Homo sapiens (human)  
PN JP 2001269182-A/3786

PD 02-OCT-2001  
PR 24-FEB-2000 JP 2000118773  
PI 26-FEB-1999 US 60/122487  
PI JEAN BAPTIST DUMAS MILNE EDWARDS, EIMERIC DUCCLAIR, JEAN YVES  
PI JORDAN  
PC C12N15/09, C07K14/435, C07K16/18, C12N1/15, C12N1/19, C12N1/21, PC  
C12N5/10,  
PC C12P21/02, C12P21/08, C12Q1/68, G06F17/30, C12N15/00, C12N5/00, PC  
G06F15/40  
CC score 7.7  
CC seq LCLGSSDSPASA/SQ  
FH key Location/Qualifiers  
FT CDS 12..347  
FT sig\_peptide 12..158.  
FEATURES  
source  
1..381 Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
ORIGIN  
Query Match 1.7%; Score 52; DB 6; Length 381;  
Best Local Similarity 100.0%; Pred. No. 5e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3071 CAAGATTGTCACATGCACTCCAGGCTGGGCAAGAGCAAGACTGTGCTC 3122  
DB 107 CAAGATTGTCACATGCACTCCAGGCTGGGCAAGAGCAAGACTGTGCTC 56  
RESULT 44  
LOCUS AX887930/c 381 bp DNA linear PAT 18-DEC-2003  
DEFINITION Sequence 3793 from Patent EP1033401.  
ACCESSION AX887930  
VERSION AX887930.1 GI:40046583  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE  
1 Dumas Milne Edwards, J.B., Ducclair, A. and Giordano, J.Y.  
AUTHORS Expressed sequence tags and encoded human proteins  
TITLE Patent: EP 1033401-A 3793 06-SEP-2000;  
JOURNAL Genet (FR)  
FEATURES  
source  
1..381 Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:9606"  
12..350  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CA893855.1"  
/db\_xref="GI:40046584"  
/translation="MHVVVVVVVVVVVVVLRGSLALLPRBCSGTILAHCSLLCLGSSD  
SPASASQVARTGMCHTQILFVFLIEGFLHIGQALBELTSGDPPASASQAGITG  
VYTPGRLSN"  
12..158  
/note="score 7.7 seq LCLGSSDSPASA/SQ"  
ORIGIN  
Query Match 1.7%; Score 52; DB 6; Length 381;  
Best Local Similarity 100.0%; Pred. No. 5e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3071 CAAGATTGTCACATGCACTCCAGGCTGGGCAAGAGCAAGACTGTGCTC 3122  
DB 107 CAAGATTGTCACATGCACTCCAGGCTGGGCAAGAGCAAGACTGTGCTC 56



RESULT 45  
AB135735/c 428 bp DNA linear STS 02-APR-2005  
LOCUS Homo sapiens DNA, STS on chromosome 1, D1S03061, sequence tagged site.  
ACCESSION AB135735  
VERSION AB135735 GI:62156316  
KEYWORDS STS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.

REFERENCE  
AUTHORS 1  
Tamiya,G., Shinya,M., Ikuta,T., Makino,S., Matsumoto,T., Mano,S., Furugaki,K., Ando,S., Nozaki,Y., Yukawa,W., Nakashige,R., Yamaguchi,D., Ishibashi,H., Yonekura,M., Nakami,Y., Takayama,S., Endo,T., Saruwatari,T., Yagura,M., Yoshikawa,Y., Fujimoto,K., Oka,A., Chiku,S., Linsen,S.E., Giphart,M.J., Bahram,S., Kulkarni,Y.V., Fukazawa,T., Hashimoto,H., Hoshina,Y., Suzuki,Y., Hotta,T., Mochida,J., Minezaki,T., Komai,K., Shiozawa,S., Taniguchi,A., Yamanaka,H., Kametani,N., Imanishi,T., Gajdosi,T. and Inoko,H.  
A Whole Genome Association Study of Rheumatoid Arthritis using 27,039 Microsatellite Markers  
JOURNAL 2 (bases 1 to 428)  
AUTHORS Tamiya,G., Makino,S., Fujimoto,K., Oka,A., Hayashi,H., Denda,A., Linsen,S.E., Ikuta,T., Shinya,M., Endo,T., Tomizawa,M., Tokubo,E., Sato,R., Takaki,A., Nagatsuka,Y., Matanabe,H., Adachi,S., Makino,Y., Nakano,S., Yamamoto,A., Yoshida,K., Okamoto,K., Yamaguchi,D., Ishibashi,H., Yonekura,M., Takayama,S., Nakami,Y., Saruwatari,T., Brand,A., van Hilten,J.A., van de Watering,L.M., Giphart,M.J., Bahram,S., Kulkarni,Y.V. and Inoko,H.  
Direct Substitution  
Submitted (12-DEC-2003) Hidetoshi Inoko, Tokai University School of Medicine, Department of Genetic Information, Bohseidai, Isehara, Kanagawa, 259-1193, Japan (E-mail:hokoketsu.tcc-u-tokai.ac.jp, Tel:81-463-93-1121, Fax:81-463-94-8884)  
Polymorphisms were confirmed by comparing with the pooled DNA typing data of 88 Dutch population or 100 Australian. When there was no significant difference between the Japanese and these Caucasians, individual typing was performed to confirm polymorphisms.

COMMENT  
Location/Qualifiers  
1..428  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/map="1p36.23"  
/note="pooled DNA of 100 Japanese unrelated individuals  
sequence tagged site D1S03061"  
1..24  
/note="5' primer: AACTGTAGAAATCTTAACGGTGAGT"  
/PCR\_conditions="denaturation 96degC 5 min, 57degC 1 min, 72degC 1 min; 40 cycles 96degC 45 sec, 57degC 45 sec, 72degC 1 min"  
80..111  
/note="sequence tagged site D1S03061"  
/rpt\_type=tandem  
/rpt\_unit="gttt"  
/complement(405..428)  
/note="3' primer: AAACATGATGCTCACTTTAGAAC"

FEATURES  
ORIGIN  
primer\_bind  
repeat\_region  
primer\_bind

Query Match 1..7%; Score 52; DB 10; Length 428;  
Best Local Similarity 100.0%; Pred. No. 5.1e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTGCTC 3122  
DB 167 CAAGATTGTGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTGCTC 116

RESULT 46  
BV640976/c 690 bp DNA linear STS 16-APR-2005  
LOCUS BV640976/c  
DEFINITION sequence tagged site.  
ACCESSION BV640976  
VERSION BV640976 GI:62668946  
KEYWORDS STS.  
SOURCE Pan troglodytes troglodytes  
ORGANISM Pan troglodytes troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Pan.

REFERENCE  
AUTHORS 1 (bases 1 to 690)  
Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and Jaffe,D.B.  
Initial Sequence of the Chimpanzee Genome and Comparison with the Human Genome  
Unpublished (2005)

TITLE  
JOURNAL  
COMMENT  
Contact: Michael C. Zody  
Broad Institute of MIT and Harvard  
320 Charles Street, Cambridge, MA 02141, USA  
Tel: 6172580933  
Fax: 6172580903  
Email: mczody@broad.mit.edu  
primer A: No sequence submitted  
primer B: No sequence submitted  
STS size: 690  
Protocol:  
23,021,928 chimpanzee whole genome shotgun reads were aligned to the Human genome NCBI Build 34 (hg16, July 2003). Chimp WGS reads were from 9 donors, including Clint (Pan troglodytes verus), 3 other Pan troglodytes verus chimps (Donald, Karlen, Yvonne), 3 Pan troglodytes troglodytes chimps (Noemie, Masuku, Clara) and 2 chimps of unknown origin (Gon, Unknown Chimp). Common names: Pan troglodytes verus is the western chimp and Pan troglodytes troglodytes is the central chimp. To be included in read must be at least 500bp in length, at least 50% of its base calls must have Phred score >= 20, at least 30% of its base calls must satisfy SNOS(30,25) (single strand NGS, the base in question has Phred score >= 30, the surrounding 10 bases in the read have Phred score >= 25), and the read must have at least 200 bp SNOS(30,25) bases. Reads not uniquely placed in the genome and read pairs whose two ends were not consistently placed were discarded. After above filtering, NGS(30,25) standard was applied to all pairs of overlapping reads to call NGS bases and SNPs. Alignments (between two reads) with less than 100 NGS bases or with SNP rate > 0.01 were discarded. To exclude alignment between two copies of a single read, comparisons between two reads that share 95% of their genome alignments (>95% bases of read A and >95% bases of read B were placed at the same locus in human genome) were discarded.

FEATURES  
ORIGIN  
location/Qualifiers  
1..690  
/organism="Pan troglodytes troglodytes"  
/mol\_type="genomic DNA"  
/sub\_species="troglodytes"  
/db\_xref="taxon:37011"  
/clone\_id="Noemie"  
<1..>690

STS

ORIGIN

Query Match 1.7%; Score 52; DB 10; Length 690;  
Best Local Similarity 100.0%; Pred. No. 5,1e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGACAACTGTCTC 3122  
|||||  
212 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGACAACTGTCTC 161

Db

RESULT 47  
BV533342 762 bp DNA linear STS 08-APR-2005  
LOCUS G591P628298RB9.T0 Clint Pan troglodytes versus STS genomic, sequence  
DEFINITION tagged site.  
ACCESSION BV533342  
VERSION BV533342.1 GI:62411126  
KEYWORDS STS.  
SOURCE Pan troglodytes versus  
ORGANISM Pan troglodytes versus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Pan.  
1 (bases 1 to 762)  
Mikkelsen,T.S., Hillier,W.L., Eichler,E.B., Zody,M.C. and  
Jaffe,D.B.  
Initial Sequence of the Chimpanzee Genome and Comparison with the  
Human Genome  
Unpublished (2005)

JOURNAL  
COMMENT

Contact: Michael C. Zody  
Broad Institute of MIT and Harvard  
320 Charles Street, Cambridge, MA 02141, USA  
Tel: 6172580933  
Fax: 6172580903  
Email: mczody@broad.mit.edu  
Primer A: No sequence submitted  
Primer B: No sequence submitted  
STS size: 762

Protocol:  
23,021,928 chimpanzee whole genome shotgun reads were aligned to  
the Human genome NCBI  
Build 34 (hg16, July 2003). Chimp WGS reads were from 9 donors,  
including Clint (Pan  
troglodytes versu), 3 other Pan troglodytes versus chimps  
(Donald,Karlien,Yvonne), 3 Pan  
troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps  
of unknown origin  
(Gon,Unknown Chimp). Common names: Pan troglodytes versus is the  
western chimp and Pan  
troglodytes troglodytes is the central chimp. To be included in  
chimpanzee SNP discovery, a  
read must be at least 500bp in length, at least 50% of its base  
calls must have Phred  
score >= 20, at least 30% of its base calls must satisfy  
SNOS(30,25)(single strand NOS, the  
base in question has Phred score >= 30, the surrounding 10 bases in  
the read have Phred  
score >= 25), and the read must have at least 200 bp SNOS(30,25)  
bases. Reads not uniquely  
placed in the genome and read pairs whose two ends were not  
consistently placed were  
discarded. After above filtering, NOS(30,25) standard was applied  
to all pairs of  
overlapping reads to call NOS bases and SNPs. Alignments (between  
two reads) with less  
than 100 NOS bases or with SNP rate > 0.01 were discarded. To  
exclude alignment between two  
copies of a single read, comparisons between two reads that share  
95% of their genome  
alignments (>=95% bases of read A and >=95% bases of read B were  
placed at the same locus

FEATURES  
source of human genome) were discarded.  
location/Qualifiers  
1..762  
/organism="Pan troglodytes versus"  
/mol\_type="genomic DNA"  
/sub\_species="versus"  
/db\_xref="taxon:37012"  
/clone\_1fb="Clint"  
<1..>762

ORIGIN  
STS

Query Match 1.7%; Score 52; DB 10; Length 762;  
Best Local Similarity 100.0%; Pred. No. 5,1e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGACAACTGTCTC 3122  
|||||  
8 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGACAACTGTCTC 59

Db

RESULT 48  
BV536183 770 bp DNA linear STS 08-APR-2005  
LOCUS G591P631495FPD4.T0 Clint Pan troglodytes versus STS genomic, sequence  
DEFINITION tagged site.  
ACCESSION BV536183  
VERSION BV536183.1 GI:62413977  
KEYWORDS STS.  
SOURCE Pan troglodytes versus  
ORGANISM Pan troglodytes versus  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Pan.  
1 (bases 1 to 770)  
Mikkelsen,T.S., Hillier,W.L., Eichler,E.B., Zody,M.C. and  
Jaffe,D.B.  
Initial Sequence of the Chimpanzee Genome and Comparison with the  
Human Genome  
Unpublished (2005)

JOURNAL  
COMMENT

Contact: Michael C. Zody  
Broad Institute of MIT and Harvard  
320 Charles Street, Cambridge, MA 02141, USA  
Tel: 6172580933  
Fax: 6172580903  
Email: mczody@broad.mit.edu  
Primer A: No sequence submitted  
Primer B: No sequence submitted  
STS size: 770

Protocol:  
23,021,928 chimpanzee whole genome shotgun reads were aligned to  
the Human genome NCBI  
Build 34 (hg16, July 2003). Chimp WGS reads were from 9 donors,  
including Clint (Pan  
troglodytes versu), 3 other Pan troglodytes versus chimps  
(Donald,Karlien,Yvonne), 3 Pan  
troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps  
of unknown origin  
(Gon,Unknown Chimp). Common names: Pan troglodytes versus is the  
western chimp and Pan  
troglodytes troglodytes is the central chimp. To be included in  
chimpanzee SNP discovery, a  
read must be at least 500bp in length, at least 50% of its base  
calls must have Phred  
score >= 20, at least 30% of its base calls must satisfy  
SNOS(30,25)(single strand NOS, the  
base in question has Phred score >= 30, the surrounding 10 bases in  
the read have Phred  
score >= 25), and the read must have at least 200 bp SNOS(30,25)  
bases. Reads not uniquely  
placed in the genome and read pairs whose two ends were not  
consistently placed were  
discarded. After above filtering, NOS(30,25) standard was applied

to all pairs of overlapping reads to call NOS bases and SNPs. Alignments (between two reads) with less than 100 NOS bases or with SNP rate > 0.01 were discarded. To exclude alignment between two copies of a single read, comparisons between two reads that share 95% of their genome alignments (>95% bases of read A and >95% bases of read B were placed at the same locus of human genome) were discarded.

## FEATURES

## source

```
1. 770
/organism="Pan troglodytes verus"
/mol_type="genomic DNA"
/sub_species="verus"
/db_xref="taxon:37012"
/clone_1b="Clint"
<1..>770
```

## ORIGIN

## STS

Query Match 1.7%; Score 52; DB 10; Length 770;  
Best Local Similarity 100.0%; Pred. No. 5.1e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCTGCGCAACAGACAGACTCTGTCTC 3122  
Db 21 CAAGATTGTGCCACTGCACTCCAGCTGCGCAACAGACAGACTCTGTCTC 72

RESULT 49  
BV588891 773 bp DNA linear STS 12-APR-2005  
LOCUS G591P638732PFA.T0 Clint Pan troglodytes verus STS genomic, sequence  
DEFINITION tagged site.  
ACCESSION BV588891  
VERSION BV588891.1 GI:62504620  
KEYWORDS STS.  
SOURCE Pan troglodytes verus  
ORGANISM Pan troglodytes verus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homniidae; Pan.  
1 (bases 1 to 773)  
Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and  
Jaffe,D.B.  
Initial Sequence of the Chimpanzee Genome and Comparison with the  
Human Genome  
Unpublished (2005)

REFERENCE  
AUTHORS Contact: Michael C. Zody  
Broad Institute of MIT and Harvard  
320 Charles Street, Cambridge, MA 02141, USA  
Tel: 6172580933  
Fax: 6172580903  
Email: mczody@broad.mit.edu  
Primer A: No sequence submitted  
Primer B: No sequence submitted  
STS size: 773  
JOURNAL  
COMMENT

Protocol: 23,021,928 chimpanzee whole genome shotgun reads were aligned to the Human genome NCBI Build 34 (hg16, July 2003). Chimp WGS reads were from 9 donors, including Clint (Pan troglodytes verus), 3 other Pan troglodytes verus chimps (Donald, Karlien, Yvonne), 3 Pan troglodytes troglodytes chimps (Noemie, Masuku, Clara) and 2 chimps of unknown origin (Gon, Unknown chimp). Common names: Pan troglodytes verus is the western chimp and Pan troglodytes troglodytes is the central chimp. To be included in chimpanzee SNP discovery, a read must be at least 500bp in length, at least 50% of its base calls must have Phred

score >= 20, at least 30% of its base calls must satisfy SNOS(30,25) (single strand NOS, the base in question has Phred score >= 30, the surrounding 10 bases in the read have Phred score >= 25), and the read must have at least 200 bp SNOS(30,25) bases. Reads not uniquely placed in the genome and read pairs whose two ends were not consistently placed were discarded. After above filtering, NOS(30,25) standard was applied to all pairs of overlapping reads to call NOS bases and SNPs. Alignments (between two reads) with less than 100 NOS bases or with SNP rate > 0.01 were discarded. To exclude alignment between two copies of a single read, comparisons between two reads that share 95% of their genome alignments (>95% bases of read A and >95% bases of read B were placed at the same locus of human genome) were discarded.

## FEATURES

## source

```
1. 773
/organism="Pan troglodytes verus"
/mol_type="genomic DNA"
/sub_species="verus"
/db_xref="taxon:37012"
/clone_1b="Clint"
<1..>773
```

## ORIGIN

## STS

Query Match 1.7%; Score 52; DB 10; Length 773;  
Best Local Similarity 100.0%; Pred. No. 5.1e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCTGCGCAACAGACAGACTCTGTCTC 3122  
Db 33 CAAGATTGTGCCACTGCACTCCAGCTGCGCAACAGACAGACTCTGTCTC 84

RESULT 50  
BV492390 775 bp DNA linear STS 07-APR-2005  
LOCUS S221P60653PFA.T0 Yvonne Pan troglodytes troglodytes STS genomic, sequence tagged site.  
DEFINITION  
ACCESSION BV492390  
VERSION BV492390.1 GI:62356060  
KEYWORDS STS.  
SOURCE Pan troglodytes troglodytes  
ORGANISM Pan troglodytes troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homniidae; Pan.  
1 (bases 1 to 775)  
Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and  
Jaffe,D.B.  
Initial Sequence of the Chimpanzee Genome and Comparison with the  
Human Genome  
Unpublished (2005)

REFERENCE  
AUTHORS Contact: Michael C. Zody  
Broad Institute of MIT and Harvard  
320 Charles Street, Cambridge, MA 02141, USA  
Tel: 6172580933  
Fax: 6172580903  
Email: mczody@broad.mit.edu  
Primer A: No sequence submitted  
Primer B: No sequence submitted  
STS size: 775  
JOURNAL  
COMMENT

Protocol: 23,021,928 chimpanzee whole genome shotgun reads were aligned to the Human genome NCBI Build 34 (hg16, July 2003). Chimp WGS reads were from 9 donors, including Clint (Pan troglodytes verus), 3 other Pan troglodytes verus chimps

(Donald,Karlén,Yvonne), 3 Pan troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps of unknown origin (Gon,Unknown Chimp). Common names: Pan troglodytes versus is the western chimp and Pan troglodytes troglodytes is the central chimp. To be included in chimpanzee SNP discovery, a read must be at least 500bp in length, at least 50% of its base calls must have Phred score >= 20, at least 30% of its base calls must satisfy SNOS(30,25) (single strand NOS, the base in question has Phred score >= 30, the surrounding 10 bases in the read have Phred score >= 25), and the read must have at least 200 bp SNOS(30,25) bases. Reads not uniquely placed in the genome and read pairs whose two ends were not consistently placed were discarded. After above filtering, NOS(30,25) standard was applied to all pairs of overlapping reads to call NOS bases and SNPs. Alignments (between two reads) with less than 100 NOS bases or with SNP rate > 0.01 were discarded. To exclude alignment between two copies of a single read, comparisons between two reads that share 95% of their genome alignments (>=95% bases of read A and >=95% bases of read B were placed at the same locus of human genome) were discarded.

## FEATURES

source

```
1..775
/organism="Pan troglodytes troglodytes"
/mol_type="genomic DNA"
/sub_species="troglodytes"
/db_xref="taxon:37011"
/clone_lib="Yvonne"
<1..>775
```

## ORIGIN

STS

Query Match 1.7%; Score 52; DB 10; Length 775;  
Best Local Similarity 100.0%; Pred. No. 5,1e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGACGATGATCCTGAGGCGAGAGTTCAGAGCCGCGGCAACAT 2940  
|||||  
Db 408 GAGGACGATGATCCTGAGGCGAGAGTTCAGAGCCGCGGCAACAT 459

RESULT 51  
BV523612/c 810 bp DNA linear STS 08-APR-2005  
DEFINITION BV523612  
G591P600782FH3.T0 Clint Pan troglodytes versus STS genomic, sequence tagged site.  
ACCESSION BV523612  
VERSION BV523612.1 GI:62401382  
KEYWORDS STS.  
SOURCE Pan troglodytes versus  
ORGANISM Pan troglodytes versus  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Pan.  
REFERENCE 1 (bases 1 to 810)  
Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and Jaffe,D.B.  
TITLE Initial Sequence of the Chimpanzee Genome and Comparison with the Human Genome  
JOURNAL Unpublished (2005)  
COMMENT Contact: Michael C. Zody  
Broad Institute of MIT and Harvard  
320 Charles Street, Cambridge, MA 02141, USA  
Tel: 6172580933  
Fax: 6172580903  
Email: mczody@broad.mit.edu

Primer A: No sequence submitted  
Primer B: No sequence submitted  
STS size: 810

## Protocol:

23,021,928 chimpanzee whole genome shotgun reads were aligned to the Human genome NCBI Build 34 (hg16, July 2003). Chimp WGS reads were from 9 donors, including Clint (Pan troglodytes versus), 3 other Pan troglodytes versus chimps (Donald,Karlén,Yvonne), 3 Pan troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps of unknown origin (Gon,Unknown Chimp). Common names: Pan troglodytes versus is the western chimp and Pan troglodytes troglodytes is the central chimp. To be included in chimpanzee SNP discovery, a read must be at least 500bp in length, at least 50% of its base calls must have Phred score >= 20, at least 30% of its base calls must satisfy SNOS(30,25) (single strand NOS, the base in question has Phred score >= 30, the surrounding 10 bases in the read have Phred score >= 25), and the read must have at least 200 bp SNOS(30,25) bases. Reads not uniquely placed in the genome and read pairs whose two ends were not consistently placed were discarded. After above filtering, NOS(30,25) standard was applied to all pairs of overlapping reads to call NOS bases and SNPs. Alignments (between two reads) with less than 100 NOS bases or with SNP rate > 0.01 were discarded. To exclude alignment between two copies of a single read, comparisons between two reads that share 95% of their genome alignments (>=95% bases of read A and >=95% bases of read B were placed at the same locus of human genome) were discarded.

## FEATURES

source

```
1..810
/organism="Pan troglodytes versus"
/mol_type="genomic DNA"
/sub_species="versus"
/db_xref="taxon:37012"
/clone_lib="Clint"
<1..>810
```

## ORIGIN

STS

Query Match 1.7%; Score 52; DB 10; Length 810;  
Best Local Similarity 100.0%; Pred. No. 5,1e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGACATGACCTCCAGGCTTGAGGCAAGAGCAAGACTCTGTCTC 3122  
|||||  
Db 392 CAAGATTGTGACATGACCTCCAGGCTTGAGGCAAGAGCAAGACTCTGTCTC 341

RESULT 52  
BV465101 915 bp DNA linear STS 06-APR-2005  
LOCUS BV465101  
DEFINITION BV465101  
G591P61938C7.T0 Clint Pan troglodytes versus STS genomic, sequence tagged site.  
ACCESSION BV465101  
VERSION BV465101.1 GI:62224529  
KEYWORDS STS.  
SOURCE Pan troglodytes versus  
ORGANISM Pan troglodytes versus  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Pan.  
REFERENCE 1 (bases 1 to 915)  
Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and Jaffe,D.B.  
TITLE Initial Sequence of the Chimpanzee Genome and Comparison with the

JOURNAL  
COMMENT

Human Genome  
Unpublished (2005)

Contact: Michael C. Zody  
Broad Institute of MIT and Harvard  
320 Charles Street, Cambridge, MA 02141, USA  
Tel: 6172580933  
Fax: 6172580903  
Email: mczody@broad.mit.edu

Primer A: No sequence submitted  
Primer B: No sequence submitted  
STS size: 915

Protocol:  
23,021,928 chimpanzee whole genome shotgun reads were aligned to the Human genome NCBI Build 34 (hg16, July 2003). Chimp WGS reads were from 9 donors, including Clint (Pan troglodytes verus), 3 other Pan troglodytes verus chimps (Donald, Karlén, Yvonne), 3 Pan troglodytes troglodytes chimps (Noemi, Masuku, Clara) and 2 chimps of unknown origin (Gon, Unknown Chimp). Common names: Pan troglodytes verus is the western chimp and Pan troglodytes troglodytes is the central chimp. To be included in chimpanzee SNP discovery, a read must be at least 500bp in length, at least 50% of its base calls must have Phred score >= 20, at least 30% of its base calls must satisfy SNOS(30,25) (single strand NOS, the base in question has Phred score >= 30, the surrounding 10 bases in the read have Phred score >= 25), and the read must have at least 200 bp SNOS(30,25) bases. Reads not uniquely placed in the genome and read pairs whose two ends were not consistently placed were discarded. After above filtering, NOS(30,25) standard was applied to all pairs of overlapping reads to call NOS bases and SNPs. Alignments (between two reads) with less than 100 NOS bases or with SNP rate > 0.01 were discarded. To exclude alignment between two copies of a single read, comparisons between two reads that share 95% of their genome alignments (>95% bases of read A and >95% bases of read B were placed at the same locus in human genome) were discarded.

## FEATURES

source

```
/organism="Pan troglodytes verus"
/mol_type="genomic DNA"
/sub_species="verus"
/db_xref="taxon:37012"
/clone_lib="Clint"
<1..>915
```

## ORIGIN

STS

Query Match 1.7%; Score 52; DB 10; Length 915;  
Best Local Similarity 100.0%; Pred. No. 5.1e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGACGATGATCAGCTGAGGACGAGATTGAGACGAGCTGAGCAACAT 2940  
|||||  
DB 651 GAGGACGATGATCAGCTGAGGACGAGATTGAGACGAGCTGAGCAACAT 702

RESULT 53  
AB014078/c 22738 bp DNA linear PRI 20-NOV-1999  
LOCUS Homo sapiens genomic DNA, chromosome 6p21.3, HLA class I region,  
DEFINITION Cosmid clone:TY1C6, complete sequence.  
ACCESSION AB014078  
VERSION AB014078.1 GI:5672590  
KEYWORDS HTG.

SOURCE  
ORGANISM

Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.

REFERENCE  
AUTHORS

Shima, T., Taniya, G., Oka, A., Takishima, N., Yamagata, T.,  
Kikkawa, E., Iwata, K., Tomizawa, M., Okuaki, N., Kuvano, Y.,  
Matanabe, K., Fukuzumi, Y., Itakura, S., Sugawara, C., Ono, A.,  
Yamazaki, M., Tashiro, H., Ando, A., Ikemura, T., Soeda, E., Kimura, M.,  
Bahram, S. and Inoko, H.  
Molecular dynamics of MHC genesis unraveled by sequence analysis of  
the 1,796,938-bp HLA class I region  
Proc. Natl. Acad. Sci. U.S.A. 96 (23), 13282-13287 (1999)

TITLE

JOURNAL

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

```
1..22738
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/map="6p21.3"
/clone="745D12 and 960H11"
/haplotype="HLA-DR4"
/sex="male"
/cell_line="Boleth"
/cell_type="Immune response cell"
/tissue_type="Blood"
/clone_lib="CEPH YAC"
/dev stage="adult"
/note="Cosmid Clone TY1C6"
```

## ORIGIN

Query Match 1.7%; Score 52; DB 8; Length 22738;  
Best Local Similarity 100.0%; Pred. No. 5.5e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGGCCACTGCAGCTCGAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
|||||  
DB 19661 CAAGATTGGCCACTGCAGCTCGAGCTGGGCAACAGAGCAAGACTGTCTC 19610

RESULT 54

LOCUS BX248096 23481 bp DNA linear PRI 28-MAY-2004  
DEFINITION Human DNA sequence from clone DASS-205E5 on chromosome 6, complete  
sequence.

ACCESSION  
VERSION

KEYWORDS  
SOURCE

ORGANISM

Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (28-MAY-2004) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humuery@sanger.ac.uk  
On May 28, 2004 this sequence version replaced gi:47776042.  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs constructed by the MHC HaploTYPE Consortium and collaborators. Further information can be found at <http://www.sanger.ac.uk/HGP/Chrs/MHC>

DASS-205E5 is from a DNA-arts S8TO human bac library VECTOR: pBelobAC11.

#### FEATURES

##### source

```
Location/Qualifiers
1..23481
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="DASS-205E5"
/clone_lib="DNA-arts-BAC-1-S8TO.1"
```

##### ORIGIN

```
Query Match      1.7%; Score 52; DB 8; Length 23481;
Best Local Similarity 100.0%; Pred.No.5.5e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3071 CAAGATTGTGCGACCTGACCTGAGCTGGGCAACAGAGCAAGCTGTCTC 3122
|||||
Db      7576 CAAGATTGTGCGACCTGACCTGAGCTGGGCAACAGAGCAAGCTGTCTC 7627
```

#### RESULT 55

AC087649/c

##### LOCUS

AC087649 25057 bp DNA linear PRI 07-MAY-2002  
Homo sapiens chromosome 17, clone RP11-46122, complete sequence.

##### ACCESSION

##### VERSION

AC087649.6 GI:204866426

##### KEYWORDS

HTG.

##### SOURCE

Homo sapiens (human)

##### ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.

##### REFERENCE

1 (bases 1 to 25057)  
Birren,B., Linton,L., Nuebaum,C. and Lander,E.  
Homo sapiens chromosome 17, clone RP11-46122  
unpublished

##### AUTHORS

2 (bases 1 to 25057)  
Birren,B., Linton,L., Nuebaum,C., Lander,E., Allen,N., Anderson,S.,  
Barna,N., Bastien,V., Boguslavsky,L., Boughgalter,B., Brown,A.,  
Camarat,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,  
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,  
Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,  
Gardyna,S., Ginde,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
Jones,C., Karatas,A., Laroque,K., Lamazares,R., Landers,T.,  
Lehoczky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,  
Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K.,

McPeters,R., Meldrim,J., Meneus,L., Mhova,T., Mlenga,V.,  
Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,  
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,  
Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R.,  
Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M.,  
Roy,A., Santos,R., Schauer,S., Schuppach,R., Seaman,S., Severy,P.,  
Sounez,C., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Straus,N., Subramanian,A., Talams,J., Testaye,S., Theodore,J.,  
Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A.,  
Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J.,  
Zembek,L., Zimmer,A. and Zody,M.

#### TITLE

Submitted (15-JAN-2001) Whitehead Institute/MIT Center for Genome

#### JOURNAL

Research, 320 Charles Street, Cambridge, MA 02141, USA

#### REFERENCE

3 (bases 1 to 25057)

Birren,B., Linton,L., Nuebaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,  
Boughgalter,B., Brown,A., Camarat,J., Campopiano,A., Chang,J.,  
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A.,  
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
Kamat,A., Karatas,A., Kelle,C., Laroque,K., Lamazares,R.,  
Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,  
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,  
Mhova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,  
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuppach,R.,  
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Straus,N., Subramanian,A., Talams,J., Testaye,S., Theodore,J.,  
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

#### TITLE

Submitted (06-MAY-2002) Whitehead Institute/MIT Center for Genome

#### JOURNAL

Research, 320 Charles Street, Cambridge, MA 02141, USA

#### REFERENCE

4 (bases 1 to 25057)

Birren,B., Linton,L., Nuebaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L.,  
Boughgalter,B., Brown,A., Camarat,J., Campopiano,A., Chang,J.,  
Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A.,  
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
Kamat,A., Karatas,A., Kelle,C., Laroque,K., Lamazares,R.,  
Landers,T., Lehoczky,J., Levine,R., Lindblad-Toh,K., Liu,G.,  
Maclean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C.,  
McCarthy,M., McEwan,P., McKernan,K., Meldrim,J., Meneus,L.,  
Mhova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R.,  
Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V.,  
Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P.,  
Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schuppach,R.,  
Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Straus,N., Subramanian,A., Talams,J., Testaye,S., Theodore,J.,  
Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

#### TITLE

Submitted (07-MAY-2002) Whitehead Institute/MIT Center for Genome

#### JOURNAL

Research, 320 Charles Street, Cambridge, MA 02141, USA

#### COMMENT

On May 7, 2002 this sequence version replaced gi:20455562.  
All repeats were identified using RepeatMasker:  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

Smit, A.F.A. & Green, P. (1996-1997)  
Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR





Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [veg@sanger.ac.uk](mailto:veg@sanger.ac.uk)

-----  
Draft Sequence Produced by Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA  
<http://genome.wustl.edu/gsc/index.shtml>

This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one subclone; and the assembly was confirmed by restriction digest,  
except on the rare occasion of the clone being a YAC.

## FEATURES

## source

1..25242  
Location/Qualifiers

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/clone\_id="RP11-308L13"  
/clone\_1ib="RPCT-11.2"  
2000  
/note="Clone\_right\_end: RP11-103C3"  
23243  
/note="Clone\_left\_end: RP11-191N8"

## misc\_feature

23243  
/note="Clone\_left\_end: RP11-191N8"

## ORIGIN

## Query Match

Best Local Similarity 1.7%; Score 52; DB 8; Length 25242;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGGCCACTGCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122

Db 3472 CAAGATTGGCCACTGCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3523

## RESULT 57

## CR36916

## LOCUS

CR36916 25972 bp DNA linear PRI 03-MAR-2005

## DEFINITION

Human DNA sequence from clone DAMC-259M15 on chromosome 6, complete  
sequence.

## ACCESSION

## CR36916

## VERSION

## CR36916.2

## KEYWORDS

## HTG.

## SOURCE

## Homo sapiens

## ORGANISM

## Homo sapiens

## Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

## Homnidae; Homo.

## 1 (bases 1 to 25972)

## Direct Submission

## Submitted (03-MAR-2005)

## Wellcome Trust Sanger Institute, Hinxton,

## Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

[humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk)

## On Mar 3, 2005 this sequence version replaced gi:604579377.

## ----- Genome Center

## Center: Wellcome Trust Sanger Institute

## Center code: SC

Web site: <http://www.sanger.ac.uk>Contact: [humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk)

## -----

During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate

chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest, except on the rare  
occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em; EMBL; SW; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information  
on the WORMPEP database can be found at  
[http://www.sanger.ac.uk/projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/projects/C_elegans/wormpep) This sequence  
was generated from part of bacterial clone contigs constructed by  
the MHC HaploType Consortium and collaborators. Further information  
can be found at  
<http://www.sanger.ac.uk/HGP/Chr6/MHC>  
DAMC-259M15 is from the DNA-Arts.org BAC library MCF.1 VECTOR:  
pBelobAC11.

## FEATURES

## source

1..25972  
Location/Qualifiers

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone\_id="DAMC-259M15"  
/clone\_1ib="DNA-Arts.org BAC library MCF.1"

## ORIGIN

## Query Match

Best Local Similarity 1.7%; Score 52; DB 8; Length 25972;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGGCCACTGCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122

Db 24111 CAAGATTGGCCACTGCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 24162

## RESULT 58

## CR382280

## LOCUS

CR382280 28276 bp DNA linear PRI 08-JUN-2004

## DEFINITION

Human DNA sequence from clone DAMA-277114 on chromosome 6, complete  
sequence.

## ACCESSION

## CR382280

## VERSION

## CR382280.5

## KEYWORDS

## HTG.

## SOURCE

## Homo sapiens

## ORGANISM

## Homo sapiens

## Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

## Homnidae; Homo.

## 1 (bases 1 to 28276)

## Direct Submission

## Submitted (08-JUN-2004)

## Wellcome Trust Sanger Institute, Hinxton,

## Cambridgeshire, CB10 1SA, UK. E-mail enquiries:

[humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk)

## On Jun 8, 2004 this sequence version replaced gi:47604309.

## ----- Genome Center

## Center: Wellcome Trust Sanger Institute

## Center code: SC

Web site: <http://www.sanger.ac.uk>Contact: [humquerry@sanger.ac.uk](mailto:humquerry@sanger.ac.uk)

## -----

During sequence assembly data is compared from overlapping clones.  
Where differences are found these are annotated as variations  
together with a note of the overlapping clone name. Note that the  
variation annotation may not be found in the sequence submission  
corresponding to the overlapping clone, as we submit sequences with  
only a small overlap as described above.  
This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least

one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs constructed by the MHC HaploType Consortium and collaborators. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6/MHC> DAMA-277114 is from the DNA-Arts human BAC library MANN.1 VECTOR: pBelOBAcl1

## FEATURES

source

Location/Qualifiers  
1..28276  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="DAMA-277114"  
/clone\_lib="DNA-Arts BAC library MANN.1"

## ORIGIN

Query Match 1.7%; Score 52; DB 8; Length 28276;  
Best Local Similarity 100.0%; Pred No. 5.5e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAGAGTTGTCACCTGCTCCAGCTGGGCAACAGAGCAAGCTGTCTC 3122  
|||||  
Dd 26981 CAGAGTTGTCACCTGCTCCAGCTGGGCAACAGAGCAAGCTGTCTC 27032

RESULT 59  
LOCUS CR391992 32906 bp DNA linear PRI 19-MAY-2005  
DEFINITION Human DNA sequence from clone RPI-308B4 on chromosome 1 Contains a novel gene (FLJ13171) and a DnaJ (Hsp40) homolog subfamily C member 8 (DnaJc8) pseudogene, complete sequence.

ACCESSION CR391992  
VERSION CR391992.5 GI:47232524  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 32906)  
Direct Submission  
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: [vega@sanger.ac.uk](mailto:vega@sanger.ac.uk)  
Clone requests: [clonerequest@sanger.ac.uk](mailto:clonerequest@sanger.ac.uk)  
On May 14, 2004 this sequence version replaced gi:7058867.  
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1>  
RPI-308B4 is from the library RPICT-1 constructed by the group of Plietzer de Jong. For further details see <http://www.chori.org/bacpac/home.htm>  
VECTOR: pCYPAC2

## COMMENT

----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [vega@sanger.ac.uk](mailto:vega@sanger.ac.uk)  
-----

## FEATURES

source

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

misc\_feature

gene

CDS

gene

mRNA

gene

mRNA

gene

mRNA

Location/Qualifiers  
1..32906  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="1"  
/clone="RPI-308B4"  
/clone\_lib="RPICT-1"  
1  
/note="Clone left end: RPI-308B4"  
3893..4306  
/locus\_tag="RPI-308B4.1-001"  
/pseudo  
3893..4306  
/locus\_tag="RPI-308B4.1-001"  
/note="match: protein: AAH66222 Sw:O75937 Tr:AAH57191  
Tr:O70149 Tr:O8C2M6 Tr:O8N425 Tr:O9VCH9"  
/pseudo  
/product="DnaJ (Hsp40) homolog, subfamily C member 8 (DnaJc8) pseudogene"  
join(8205..8232,9458..9631,10585..10665,16058..16222,16755..17141,23928..24525,26481..26665,30825..30978,AL840643.1:18638..8693,AL840643.1:9242..9362,AL840643.1:10419..10480,AL840643.1:11118..11255,AL840643.1:12483..12559,AL840643.1:13698..13790,AL840643.1:16002..16624)  
/gene="RPI-442N24.A.1"  
/locus\_tag="RPI-442N24.A.1-002"  
join(8205..8232,9458..9631,10585..10665,16058..16222,16755..17141,23928..24525,26481..26665,30825..30978,AL840643.1:18638..8693,AL840643.1:9242..9362,AL840643.1:10419..10480,AL840643.1:11118..11255,AL840643.1:12483..12559,AL840643.1:13698..13790,AL840643.1:16002..16624)  
/gene="RPI-442N24.A.1"  
/locus\_tag="RPI-442N24.A.1-002"  
/product="novel protein"  
/note="match: ESTs: Em:BC068508.1  
match: CDNA: Em:BC068508.1  
join(9458..9631,10585..10665,16058..16222,16755..17141,23928..24525,26481..26665,30825..30978,AL840643.1:18638..8693,AL840643.1:10419..10480,AL840643.1:11118..11255,AL840643.1:12483..12559,AL840643.1:13698..13790,AL840643.1:16002..16624)  
/gene="RPI-442N24.A.1"  
/locus\_tag="RPI-442N24.A.1-001"  
join(9458..9631,10585..10665,16058..16222,16755..17141,23928..24525,26481..26665,30825..30978,AL840643.1:18638..8693,AL840643.1:10419..10480,AL840643.1:11118..11255,AL840643.1:12483..12559,AL840643.1:13698..13790,AL840643.1:16002..16624)  
/gene="RPI-442N24.A.1"  
/locus\_tag="RPI-442N24.A.1-001"  
/product="novel protein"  
/note="match: ESTs: Em:AU131111.1 Em:AU134282.1  
Em:BF94049.1 Em:BG170714.1 Em:AF130081.1 Em:AK023233.1 Em:AK025436.1  
Em:BC021247.2 Em:BC029266.1  
join(9458..9631,10585..10665,16058..16222,16755..17141,23928..24525,26481..26665,30825..30978,AL840643.1:18638..8693,AL840643.1:10419..10480,AL840643.1:11118..11255,AL840643.1:12483..12559,AL840643.1:13698..13790,AL840643.1:16002..16624)  
/gene="RPI-442N24.A.1"  
/locus\_tag="RPI-442N24.A.1-003"  
join(9458..9631,10585..10665,16058..16222,16755..17141,23928..24525,26481..26665,30825..30978,AL840643.1:18638..8693,AL840643.1:10419..10480,AL840643.1:11118..11255,AL840643.1:12483..12559,AL840643.1:13698..13790,AL840643.1:16002..16624)  
/gene="RPI-442N24.A.1"  
/locus\_tag="RPI-442N24.A.1-003"  
/product="novel protein"

CDS

/note="match: cDNAs: Em:BC016152.1"  
 join(9490..9631,10585..10665,16058..16222,16755..17141,  
 23928..24525,26483..26665,30825..30978,  
 AL840643.1:8638..8693,AL840643.1:10413..10480,  
 AL840643.1:11118..11255,AL840643.1:12483..12559,  
 AL840643.1:16002..16017)  
 /gene="RP11-442N24\_A.1"  
 /locus\_tag="RP11-442N24\_A.1"  
 /standard\_name="OTTHUMP0000003548"  
 /note="match: proteins: Sw:O75167 Tr:Q91Z21 Tr:Q9N384  
 Tr:Q91VK6 Tr:Q9H395 Tr:Q9H6X0 Tr:Q9H8W6"  
 /codon\_start=1  
 /product="novel protein"  
 /protein\_id="CAI20111.1"  
 /db\_xref="GI:56208555"  
 /db\_xref="InterPro:IPR004018"  
 /db\_xref="UniProt/TREMBL:O8N84"  
 /translation="MTLDSEVAGDTPPKRKSGFGFKIPEKMKRKSSDKPKE  
 TSEVLERKISMRKPRELVKRGVLEDPQSGSDPKSPDAMLKNGHTTPIGNARSS  
 PVQVEEPVRLASIRKAIPEPLKRLSGTSGOPNSASVBNPKPPLPKRPLS  
 SSBASBEGAKDNTSGGTARPTISITTAAPATTAATSLAKTVNLSTPAPRTLS  
 PAAPASTNTATPSTLHWPAKOPPIPEPKPARNSNPVIALSQAINSGTLLKSP  
 PLPKRGIPSTVPTLESAAITTKTPSDERKSTCSMGSELPMISPSRSPPLPTH  
 IPBPRTPTPPPAKTFQVVEIPEPDLHQBIPQSDQKKEPKRIIDONGEPHI  
 PRLPLPLPHIRIQALTSPLMTPLIEGSHRASHLPENSDFSDSTLGRSLIP  
 ITIEMUKVPDDEEBEBCQTPSTSEMTPTSVIPKLPQCLREBEKESDSEGPVQY  
 RDBEDDESYSQSLANKVKRKTJLANKLWHRSEPLNLSWPCSKSEKWEIRHIG  
 NTLIRLSQRPTELEBQRNIIQPKNEARQARKIKRLTRKLSQRTVALLARK  
 ILRFNEYEVETDAQDVRADRPMTVLTLPADRAAIRKELNEFKSSMEVHEESKHFTR  
 YRP"

misc\_feature  
 30907  
 /gene="RP11-442N24\_A.1"  
 /locus\_tag="RP11-442N24\_A.1-001"  
 /note="Clone\_left\_end: RP11-442N24"

ORIGIN

Query Match 1.7%; Score 52; DB 8; Length 32906;  
 Best Local Similarity 100.0%; Pred. No. 5.5e-16;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAGATTGTGCACCTGACCTCCAGCTGGGCAACAGAGCAAGACTCTGCTC 3122  
 |||||  
 DB 10951 CAGAGTTGTGCCACTGCTCCAGCTGGGCAACAGAGCAAGACTCTGCTC 10900

RESULT 60  
 AC093235/c 38173 bp DNA linear PRI 29-SEP-2001  
 LOCUS Homo sapiens chromosome 19 clone L1NLR-267A4, complete sequence.  
 DEFINITION AC093235  
 AC093235 GI:15808549  
 VERSION AC093235.2  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 1 (bases 1 to 38173)  
 DOE Joint Genome Institute and Stanford Human Genome Center.  
 TITLE Direct Submission  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 38173)  
 DOE Joint Genome Institute.  
 TITLE Direct Submission  
 JOURNAL Unpublished  
 REFERENCE 3 (bases 1 to 38173)  
 Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 AUTHORs DOE Joint Genome Institute and Stanford Human Genome Center.  
 TITLES Direct Submission  
 JOURNAL Submitted (29-SEP-2001) DOE Joint Genome Institute, 2800 Mitchell  
 Drive, Walnut Creek, CA 94598, USA  
 COMMENT On Sep 29, 2001 this sequence version replaced gi:15193369.  
 Draft Sequence Produced by DOE Joint Genome Institute

www.jgi.doe.gov  
 Finishing Completed at Stanford Human Genome Center  
 www.shgc.stanford.edu  
 Quality: Phrap Quality >=40 99.9% of Sequence;  
 Estimated Total Number of Errors is 0.  
 Location/Qualifiers  
 1. 38173  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="19"  
 /clone="L1NLR-267A4"

ORIGIN

Query Match 1.7%; Score 52; DB 8; Length 38173;  
 Best Local Similarity 100.0%; Pred. No. 5.5e-16;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGAGGTGATGATCCTGAGGCGGAGATTGAGACGACCTGGCCACAT 2940  
 |||||  
 DB 19282 GAGGAGGTGATGATCCTGAGGCGGAGATTGAGACGACCTGGCCACAT 19231

RESULT 61  
 HSN121B8 39372 bp DNA linear PRI 18-MAY-2005  
 LOCUS Human DNA sequence from clone LL22NC03-121B8 on chromosome 22,  
 complete sequence.  
 DEFINITION HSN121B8  
 HSN121B8 278421  
 ACCESSION 278421.1 GI:1495467  
 VERSION 278421.1  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 1 (bases 1 to 39372)  
 Kettlborough, R.  
 DIRECT SUBMISSION  
 Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk  
 Clone requests: clonerequest@sanger.ac.uk  
 The following abbreviations are used to associate primary accession  
 numbers given in the feature table with their source databases:  
 Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormBase; Information  
 on the WormBase database can be found at  
 http://www.sanger.ac.uk/Projects/C\_elegans/wormpep  
 LL22NC03-121B8  
 is from the human chromosome 22-specific cosmid library LL22NC03,  
 constructed at the Biomedical Sciences Division, Lawrence Livermore  
 National Laboratory, Livermore, CA 94550 under the auspices of the  
 National Laboratory Gene Library Project sponsored by the US  
 Department of Energy. The source of the flow sorted chromosomes  
 was a human/hamster hybrid containing chromosomes Y, 22 and 9.  
 VECTOR: lambda1816  
 This sequence was generated from part of bacterial clone contigs of  
 human chromosome 22, constructed by the Sanger Centre Chromosome 22  
 Mapping Group. Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chr22  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: vegas@sanger.ac.uk  
 -----  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >=  
 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one subclone; and the assembly was confirmed by restriction digest,  
 except on the rare occasion of the clone being a YAC.  
 Location/Qualifiers  
 1. 39372

FEATURES  
 source

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="22"
/clone="LL22NC03-121B8"
/clone_1fb="LL22NC03"
/complement(join(5697..5931,6975..7281))
/locus_tag="LL22NC03-121B8.2-002"
/locus_tag(join(5697..5931,6975..7281))
/locus_tag="LL22NC03-121B8.2-002"
/locus_tag(join(7219..8044,8241..8279,8810..8929,9409..9514))
/locus_tag="LL22NC03-121B8.1-001"
/locus_tag(join(7219..8044,8241..8279,8810..8929,9409..9514))
/locus_tag="LL22NC03-121B8.1-001"
/locus_tag(join(7598..8044,8241..8279,8810..8866))
/locus_tag="LL22NC03-121B8.1-001"
/standard_name="OTTHUMP0000028824"
/codon_start=1
/protein_id="CA118780.1"
/db_xref="GI:56208181"
/db_xref="UniProt/TREMBL:O5R3S2"
/translation="MPPCTRHGAASAAVSPGNHPHAKRRKGSILPDCVQHGRDT
DKHIWQRHGGPMLELOQLCVTELPNBDGMPRTVCPHVGAMEDQGLTPSGF
PHEASILVGSMAKHGRCMGSSVDFSOAARPLAQAQENRNSILEQKMPSTSRPT
AQRWTLRDMLTRMLKYVLS"

```

## ORIGIN

```

Query Match      1.7%; Score 52; DB 8; Length 39372;
Best Local Similarity 100.0%; Pred. No. 5.5e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3071 CAAGATTGGCCACTGCACCTCCAGCTGGGCAACAGACAAGACTGTCTC 3122
DB      15817 CAAGATTGGCCACTGCACCTCCAGCTGGGCAACAGACAAGACTGTCTC 15868

```

```

RESULT 62
HSN121B8/c      39372 bp      DNA      linear      PRI 18-MAY-2005
LOCUS      HSN121B8
DEFINITION      Human DNA sequence from clone LL22NC03-121B8 on chromosome 22,
complete sequence.
ACCESSION      278421
VERSION      278421.1 GI:1495467
KEYWORDS
SOURCE      HTG.
ORGANISM      Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 39372)
Kettleborough/R.
Direct Submission
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
Clone requests: clonerequests@sanger.ac.uk
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Emi; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep
LL22NC03-121B8
is from the human chromosome 22-specific cosmid library LL22NC03,
constructed at the Biomedical Sciences Division, Lawrence Livermore
National Laboratory, Livermore, CA 94550 under the auspices of the
National Laboratory Gene Library Project sponsored by the US
Department of Energy. The source of the flow sorted chromosomes
was a human/hamster hybrid containing chromosomes Y, 22 and 9.
VECTOR: lawrisc16
This sequence was generated from part of bacterial clone contigs of
human chromosome 22, constructed by the Sanger Centre Chromosome 22
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr22
----- Genome Center
Center: Wellcome Trust Sanger Institute

```

```

Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
Location/Qualifiers
1..39372
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="22"
/clone="LL22NC03-121B8"
/clone_1fb="LL22NC03"
/complement(join(5697..5931,6975..7281))
/locus_tag="LL22NC03-121B8.2-002"
/locus_tag(join(5697..5931,6975..7281))
/locus_tag="LL22NC03-121B8.2-002"
/locus_tag(join(7219..8044,8241..8279,8810..8929,9409..9514))
/locus_tag="LL22NC03-121B8.1-001"
/locus_tag(join(7219..8044,8241..8279,8810..8929,9409..9514))
/locus_tag="LL22NC03-121B8.1-001"
/locus_tag(join(7598..8044,8241..8279,8810..8866))
/locus_tag="LL22NC03-121B8.1-001"
/standard_name="OTTHUMP0000028824"
/codon_start=1
/protein_id="CA118780.1"
/db_xref="GI:56208181"
/db_xref="UniProt/TREMBL:O5R3S2"
/translation="MPPCTRHGAASAAVSPGNHPHAKRRKGSILPDCVQHGRDT
DKHIWQRHGGPMLELOQLCVTELPNBDGMPRTVCPHVGAMEDQGLTPSGF
PHEASILVGSMAKHGRCMGSSVDFSOAARPLAQAQENRNSILEQKMPSTSRPT
AQRWTLRDMLTRMLKYVLS"

```

## FEATURES

```

source      1..39372
gene
mRNA
gene
CDS
mRNA

```

## ORIGIN

```

Query Match      1.7%; Score 52; DB 8; Length 39372;
Best Local Similarity 100.0%; Pred. No. 5.5e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3071 CAAGATTGGCCACTGCACCTCCAGCTGGGCAACAGACAAGACTGTCTC 3122
DB      17381 CAAGATTGGCCACTGCACCTCCAGCTGGGCAACAGACAAGACTGTCTC 17330

```

```

RESULT 63
AL445257/c      39596 bp      DNA      linear      HTG 13-JUN-2001
LOCUS      AL445257
DEFINITION      Homo sapiens chromosome 1 clone RP5-1175N1, 14 unordered pieces.
ACCESSION      AL445257
VERSION      AL445257.1 GI:10716509
KEYWORDS      HTG: HTGS_PHASE1; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1
Burton J.
Direct Submission
Submitted (12-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information

```

## COMMENT

```

REFERENCE
AUTHORS      Burton J.
TITLE      Direct Submission
JOURNAL
COMMENT
Center: Sanger Centre
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information

```

```

Center project name: d1175N1
----- Summary Statistics -----
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08753; 100% of reads
Chemistry: dye-terminator Big Dye; 100% of reads
Consensus quality: 28716 bases at least Q40
Consensus quality: 32676 bases at least Q30
Consensus quality: 35450 bases at least Q20
Insert size: 38296; sum-of-contigs
Insert size: 138118; 12.9% error; agarose-fp
Quality coverage: 1.51x in Q20 bases; sum-of-contigs quality
coverage: 0.97x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
2942: contig of 2942 bp in length
2943 3042: gap of 100 bp
3043 5573: contig of 2531 bp in length
5574 5674: gap of 100 bp
5674 9075: contig of 3402 bp in length
9076 9176: gap of 100 bp
9176 11187: contig of 2012 bp in length
11188 11287: gap of 100 bp
11288 14656: contig of 3369 bp in length
14657 14757: gap of 100 bp
14757 17428: contig of 2672 bp in length
17429 17528: gap of 100 bp
17529 19964: contig of 2436 bp in length
19965 20064: gap of 100 bp
20065 24184: contig of 4120 bp in length
24185 24285: gap of 100 bp
24285 26468: contig of 2184 bp in length
26469 26568: gap of 100 bp
26569 29667: contig of 3099 bp in length
29668 29767: gap of 100 bp
29768 31926: contig of 2159 bp in length
31927 32026: gap of 100 bp
32027 34705: contig of 2679 bp in length
34706 34805: gap of 100 bp
34806 37009: contig of 2204 bp in length
37010 39596: contig of 2487 bp in length.
37110
Location/Qualifiers
1..39596
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP5-1175N1"
/clone_1fb="RP5-1"
1..2942
/contig="assembly_fragment:00018"
fragment_chain:1"
3043..5573
/contig="assembly_fragment:00371"
fragment_chain:1"
5674..9075
/contig="assembly_fragment:00430"
fragment_chain:2"
9176..11187
/contig="assembly_fragment:00095"
fragment_chain:2"
11288..14656
/contig="assembly_fragment:00113"
14757..17428
/contig="assembly_fragment:00120"
17529..19964

```

```

misc_feature /note="assembly_fragment:00198"
20065..24184
/note="assembly_fragment:00245"
24285..26468
/note="assembly_fragment:00304"
26569..29667
/note="assembly_fragment:00328"
29768..31926
/note="assembly_fragment:00481"
32027..34705
/note="assembly_fragment:00489"
34806..37009
/note="assembly_fragment:00517"
37110..39596
/note="assembly_fragment:00532"

ORIGIN
Query Match 1.7%; Score 52; DB 14; Length 39596;
Best Local Similarity 100.0%; Pred. No. 5.5e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGTTGTGCACCTGCCTCAGCTGCGCAACAGCAAGACCTGTCTC 3122
DB 12668 CAAGTTGTGCACCTGCCTCAGCTGCGCAACAGCAAGACCTGTCTC 12617

RESULT 64
AL359082
LOCUS
DEFINITION
Human DNA sequence from clone RP4-576K7 on chromosome 1 Containing part of the FRAP1 (FK506 binding protein 12-rapamycin associated protein 1) gene. ESTs, an STS and GSSs, complete sequence.
AL359082 GI:14329962
VERSION
AL359082.16 GI:14329962
KEYWORDS
HTG; binding protein; FRAP1.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 41372)
REFERENCE
HALL, R.
AUTHORS
JOURNALS
Direct Submission
Submitted (11-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequests@sanger.ac.uk
On Jun 8, 2001 this sequence version replaced gi:13445394.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. The following
abbreviations are used to associate primary accession numbers given
in the feature table with their source databases: Em, EMBL; Sw,
SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP
database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr1
RP4-576K7 is from the library RP4-4 constructed by the group of
Pleier de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pCYPAC2

```



IMPORTANT: This sequence is not the entire insert of clone RP4-576K7. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap. The true left end of clone RP4-576K7 is at 1 in this sequence. The true left end of clone RP4-635B18 is at 41273 in this sequence.

## FEATURES

SOURCE

```
1. 41372
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP4-576K7"
/clone_id="RP4-576K7"
join(<401..540,7465..7565,10902..11017,12648..12725,
12938..13045,18036..18161,18259..18390,23230..23345,
24502..24619,26922..27170,27861..27961,28748..28844,
29147..29245,29573..29695,29893..30075,30555..30689,
30903..31077,33066..33201,35573..35720,36331..36453,
40608..>40690)
/gene="FRAP1"
join(<401..540,7465..7565,10902..11017,12648..12725,
12938..13045,18036..18161,18259..18390,23230..23345,
24502..24619,26922..27170,27861..27961,28748..28844,
29147..29245,29573..29695,29893..30075,30555..30689,
30903..31077,33066..33201,35573..35720,36331..36453,
40608..>40690)
/gene="FRAP1"
/note="match: CDNA: Em:U11681 Em:U8966 Em:AF152838
Em:L34075 Em:L37085 Em:AF046004 Em:AK024393 Em:L35478
match: ESTs: Em:BF689449 Em:BG120811 Em:BG174145
Em:BE560006 Em:BE560852 Em:AL120924 Em:BF782990
Em:BG342777 Em:BE933597 Em:BE871879 Em:A1736514
Em:BE039490 Em:AY594302 Em:U135855 Em:BE911828
Em:BE010022 Em:BF118061 Em:R11912 Em:BE153620 Em:BE921354
Em:R11865 Em:BG421771 Em:W26223 Em:BF21191 Em:R60442
Em:R55558 Em:AA497757 Em:A1614270 Em:BE514616 Em:BE720437
Em:T75267 Em:BG250647 Em:BF18028 Em:BF782326 Em:BE236792
Em:BG088519 Em:BE835098 Em:F08162 Em:BE892046 Em:T05942
Em:AK500990 Em:AY594303 Em:AL079589 Em:BE856122
Em:BF920810 Em:BF760012 Em:BF508672
match: proteins: Tr:OJULN9 Sw:P42346 Tr:O9Y413 Sw:P42345
Sw:P32600 Tr:O9Y7K2 Tr:O14356 Tr:O9VK45 Tr:O9V4189
Tr:O9LPM4 Tr:O94188 Sw:P35169 Tr:O01438 Tr:O9Y130
Tr:O76222"
/codon_start=1
/evidence=not_experimental
/product="d576K7.1 (FK506 binding protein 12-ramapycin
associated protein 1)"
/protein_id="CAC42395.1"
/db_xref="GI:14530309"
/db_xref="GOA:O960M8"
/db_xref="InterPro:IPR000403"
/db_xref="InterPro:IPR003151"
/db_xref="InterPro:IPR006025"
/db_xref="UniProt/TREMBL:O960M8"
/translation="RQAATWYKLEHEDALVAADKADTNDKDDDELMGMRCEAL
GEMGDIHQCCCKATLVNDETQAKARAAAAMKLGQDSHEBTCTIPDTHGAP
TRAVIALHODFSLAQCCIDKARDLDELTMAESSTSPRAAGAVSGMSELSEV
QYKLVBERREIROIWMERLQGCORIVEDMOKIIVNRSLVSPHGDMDRTWKYSLG
KSGRLAAKTLVLLGVDPKRDHPPLTPVQTVAYMKMGSAKRIAPFQMHQ
FQVQQAQAHALATEDDOOHOKELHAKRCFLKGEMLNQGINSSTIPVLYQS
AATEDRSYKAMAMAVNFEAVLVHKGONARBEKKKLHSAANTTNATTA
AATTTDSYKNSSEASSTNSPTSPLOKVTEDSKITLMTTVPAVQCFPSIS
LSKGNLDQTLVLTLPDYGMPVNEALVGVAFIQTWLYQVTPQILADIPRP
LVGRILHQLTLDIGRYHQAALYPLTVASKSTTARNAHNVLPILNCEHSTLAQQA
MMVSEELIRVALIEMHMEHGLEASRPVFGERNVKGHFEVPLAHMEHSPQTLKE
TSFOAYGDLMEAOBQCKWKSQNVCLTOAMLYHYFRIYKOLPOLTSLEIYF
VSPKLMCRDLAIVPGTVDPQPIRIIOSTAPSLVTSORPKRLITMSNGEIFY
FLKGHEDRODERVMQPLGVYNTILANDPSTLRKLSIQRNAVPLSTNSGLIWPV
HCDTIALIRDRREKKILNTEHRIMLMADVDYHLLQVVEFAHVNNTACDDL
AKLMLKSPSSSEWFDRTNRYTRSLAVSMGVYILGLD"
repeat_region
942..1116
/note="112 repeat: matches 2514..2688 of consensus"
```

```
repeat_region 1403..1711
/note="AluX repeat: matches 1..310 of consensus"
repeat_region 1796..1919
/note="L2 repeat: matches 2588..2709 of consensus"
repeat_region 2065..2274
/note="MER39b repeat: matches 193..386 of consensus"
repeat_region 2637..2828
/note="AluU repeat: matches 1..190 of consensus"
repeat_region 2829..2989
/note="FAM repeat: matches 1..174 of consensus"
repeat_region 2990..3046
/note="AluU repeat: matches 85..310 of consensus"
repeat_region 3302..3399
/note="AluU/FAM repeat: matches 2..99 of consensus"
repeat_region 3493..3528
/note="4 copies 9 mer tgcgtgtg 94% conserved"
repeat_region 3550..3592
/note="MER3 repeat: matches 169..209 of consensus"
repeat_region 3593..3904
/note="AluX repeat: matches 1..312 of consensus"
repeat_region 3905..4063
/note="MER3 repeat: matches 1..169 of consensus"
repeat_region 5278..5341
/note="U6 repeat: matches 1..65 of consensus"
repeat_region 6021..6203
/note="MER5A repeat: matches 1..189 of consensus"
repeat_region 6653..7159
/note="match: GSS: Em:A0686069"
repeat_region 7139..7569
/note="match: GSS: Em:A080311"
repeat_region 7462..7569
/note="match: GSS: Em:A0380853"
repeat_region 8300
/note="MER21B repeat: matches 39..481 of consensus"
repeat_region 8328..8426
/note="MER21B repeat: matches 568..669 of consensus"
repeat_region 8613..8759
/note="FRAP1"
repeat_region 8757..8963
/note="match: GSS: Em:A0553575"
repeat_region 8747
/note="AluU repeat: matches 2..122 of consensus"
repeat_region 8749..9044
/note="match: GSS: Em:A0531228"
repeat_region 8750..9015
/note="match: GSS: Em:A0019010"
repeat_region 8963
/note="FRAP1"
repeat_region 8784..9073
/note="match: GSS: Em:A0781128"
repeat_region 9073
/note="AluSg repeat: matches 1..292 of consensus"
repeat_region 8785..8897
/note="match: GSS: Em:A0515715"
repeat_region 8790..8947
/note="FRAP1"
repeat_region 9358..9583
/note="match: GSS: Em:A061904"
repeat_region 9583
/note="LTR1J repeat: matches 235..463 of consensus"
repeat_region 9626..9706
/note="L1MCS repeat: matches 7836..7916 of consensus"
repeat_region 11890..12024
/note="AluSg/x repeat: matches 1..138 of consensus"
repeat_region 12025..12313
/note="AluX repeat: matches 1..297 of consensus"
repeat_region 13341..13507
/note="MER5B repeat: matches 1..178 of consensus"
repeat_region 13761..14140
/note="match: GSS: Em:A0089815"
repeat_region 13852..14185
/note="match: GSS: Em:A02030613"
repeat_region 14226..14687
/note="FRAP1"
repeat_region 14226..14655
/note="match: GSS: Em:A0760056"
```

```

/gene="FRAP1"
/nc="match: GSS: Em:AQ186184"
repeat_region 15097..15396
/nc="AluX repeat: matches 1..292 of consensus"
repeat_region 15397..15460
/nc="AluS repeat: matches 1..69 of consensus"
repeat_region 15663..15988
/nc="AluY repeat: matches 1..311 of consensus"
repeat_region 16025..16296

Query Match 1.7%; Score 52; DB 8; Length 41372;
Best Local Similarity 100.0%; Pred.No. 5.5e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 3071 CAAGTTGTGCACTGCACCTCGAGCTGGGCAACAGACAACTCTGTCTC 3122
Db 35426 CAAGTTGTGCACTGCACCTCGAGCTGGGCAACAGACAACTCTGTCTC 35477

RESULT 65
LOCUS AC004602 42572 bp DNA linear PRI 22-APR-1998
DEFINITION Homo sapiens chromosome 19, cosmid F23487, complete sequence.
ACCESSION AC004602
VERSION AC004602.1 GI:3075375
KEYWORDS HTG;
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 42572)
Lamerdin,J.E., McCready,P.M., Skowronski,E., Adamson,A.W.,
Burkhardt-Schultz,K., Gordon,L., Kyle,A., Ramirez,M., Stillwagen,S.,
Phan,H., Velasco,N., Garner,J., Danganan,L., Poundstone,P.,
Christensen,M., Georgescu,A., Avila,J., Liu,S., Actix,C.,
Andreise,T., Frankheim,M., Amico-Keller,G., Coefield,J., Duarte,S.,
Lucas,S., Bruce,R., Thomas,P., Quan,G., Krommiller,B., Arellano,A.,
Montgomery,M., Ow,D., Nolan,M., Trong,S., Kobayashi,A., Olsen,A.O.,
and Carrano,A.V.
Sequence analysis of a 1 Mb region in 19p13.3
Unpublished
2 (bases 1 to 42572)
Lamerdin,J.E.
Direct Submission
Submitted (22-APR-1998) Joint Genome Institute, Lawrence Livermore
National Laboratory, 7000 East Ave., Livermore, CA 94551, USA
Map and sequence oriented from p telomere to centromere. Cosmid
F23487 should overlap cosmid R3035 to the right. There is
currently an approx. 8 kb sequence gap to the left between F23487
and P1-23569. Additional chr 19 map and sequence information is
available at http://www-bio.livl.gov/genome/genome.html.
location/Qualifiers
1..42572
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="19"
/map="19p13.3 between D19S883 and MLT1"
/clone="F23487"
/cell_line="UV5HL9-5B"
/clone_lib="ML19NC02 F chromosome 19-specific cosmid
library"
/nc="Cosmid library constructed at LIVL from flow-sorted
chromosomes from hybrid UV5HL9-5B, which carries
chromosome 19 as its only human chromosome."
repeat_region complement(1..63)
/rpt_family="Alu"
repeat_region complement(68..330)
/rpt_family="Alu"
repeat_region complement(333..605)
/rpt_family="Alu"
repeat_region complement(741..771)

```

---

```

repeat_region complement(769..886)
/rpt_family="LINE2"
repeat_region 1087..1388
/rpt_family="AluX"
repeat_region complement(1389..1663)
/rpt_family="MLT1A2"
repeat_region 1668..1832
/rpt_family="THE1C"
repeat_region 1833..2129
/rpt_family="AluX"
repeat_region 2139..2343
/rpt_family="THE1C"
repeat_region complement(2344..2404)
/rpt_family="MLT1A2"
repeat_region 2405..2665
/rpt_family="AluX"
repeat_region 2677..2977
/rpt_family="AluSg"
repeat_region complement(2978..3002)
/rpt_family="AT-rich"
3029..3131
/nc="Predicted exon, program: grr1lexons_human_1.3,
frame: 2, quality: good, score: 72.000"
3166..3478
/nc="Predicted exon, program: grr1lexons_human_1.3,
frame: 1, quality: excellent, score: 91.000"
complement(3820..4131)
/rpt_family="AluSx"
complement(4727..4762)
/rpt_family="MIR"
4763..5031
/rpt_family="AluSx"
5032..5055
/rpt_family="(CNA)n"
complement(5154..5254)
/rpt_family="Alu"
5256..5545
/rpt_family="AluSx"
complement(5546..5719)
/rpt_family="Alu"
complement(6013..6320)
/rpt_family="AluSx"
complement(6339..6503)
/rpt_family="AluSg"
complement(6504..6804)
/rpt_family="AluSg"
complement(6806..6942)
/rpt_family="AluSg"
complement(6945..7248)
/rpt_family="Alu"
complement(7754..8053)
/rpt_family="AluSg"
7955..8378
/nc="PDS similarity to H80195 yu56f07.r1 Homo sapiens
cDNA clone 230149.5, similar to contains Alu repetitive
element,contains HGR repetitive element ;. Score: 788
identity: 419/428 (97%)."
complement(8082..8198)
/rpt_family="FLAM C"
complement(8272..8312)
/rpt_family="MIR"
8632..10218
/gene="CAPS"
8632..8651
/gene="CAPS"
/nc="PDS similarity to R59818 yu11c08.r1 Homo sapiens
cDNA clone 42992.5, similar to SP.A49020 A49020
CALCIUM-BINDING PROTEIN R205; (1..35); 100% identity."
8649..8831
/gene="CAPS"
/standard_name="exon"
/nc="PDS similarity to overlapping ESTs:

```

(8649..8831) AA128281 z129e12.r1 Soares pregnant uterus  
NBHPU Homo sapiens cDNA clone 503374 5' similar to  
TR:E245872 E245872 CALCYPHOSINE; (1..183); 99%  
identity.- (8693..8831) AA359592 EST66631 Fetal lung II  
Homo sapiens cDNA 5' and similar to calcyphosine;  
(1..139); 99% identity.- (8727..8831) R59818 y11c08.r1  
Homo sapiens cDNA clone 42992 5' similar to SP:A49920  
A4920 CALCIUM-BINDING PROTEIN R2D5; (36..140); 99%  
identity."  
misc\_feature  
8727..8831  
/gene="CAPS"  
/note="predicted exon, program: grail2exons\_human\_1.3,  
frame: 0, quality: excellent, score: 100.000"  
join(8749..8831,8905..9082,9282..9488,9563..9664)  
/gene="CAPS"  
/note="CALCYPHOSINE (CALCIUM BINDING PROTEIN); R2D5"  
/codon\_start=1  
/product="CAMP HUMAN"  
/protein\_id="A014484.1"  
/db\_xref="GI:3075376"  
/translation="MDADVATWEKLRPAQCLSRGASGICGLARFFQIDRDSRLSDAD  
EFRQGLAGLVLDQAEAGVCRKWRNGSGTLDEFLRLRPRMSQARAVINAAE  
AKIDRSGGVVVDLRGVSGRAHPKVRSGEWDDEVLRFLDNFDSSEKQGVTLA  
EFQDYSGVASAMTDESFVAMMTSAMQL"  
8905..9082  
/gene="CAPS"  
/standard\_name="exon"  
/note="PDS similarity to overlapping ESTs:  
(8905..9082) AA128281 z129e12.r1 Soares pregnant uterus  
NBHPU Homo sapiens cDNA clone 503374 5' similar to  
TR:E245872 E245872 CALCYPHOSINE; (184..359); 98%  
identity.- (8905..9082) R59818 y11c08.r1 Homo sapiens cDNA  
clone 42992 5' similar to SP:A49920 A4920 CALCIUM-BINDING  
PROTEIN R2D5; (141..321); 93% identity.- (8905..9013)  
AA359592 EST66631 Fetal lung II Homo sapiens cDNA 5' end  
similar to calcyphosine; (140..251); 95%  
identity.- predicted exon, program: grail2exons\_human\_1.3,  
frame: 1, quality: good, score: 71.000"  
9282..9488  
/gene="CAPS"  
/standard\_name="exon"  
/note="PDS similarity to overlapping ESTs:  
(9282..9488) AA128281 z129e12.r1 Soares pregnant  
uterus NBHPU Homo sapiens cDNA clone 503374 5' similar to  
TR:E245872 E245872 CALCYPHOSINE; (360..502); 96%  
identity."  
misc\_feature  
9523..9832  
/gene="CAPS"  
/standard\_name="exon"  
/note="PDS similarity to R88929 yp97h05.r1 Homo sapiens  
cDNA clone 195417 5' similar to SP:CAMP\_CANFA P10463  
CALCYPHOSIN; (19..359); 99% identity."  
9563..9660  
/gene="CAPS"  
/note="predicted exon, program: grail2exons\_human\_1.3,  
frame: 1, quality: excellent, score: 100.000"  
complement(9722..10218)  
/note="PDS similarity to overlapping ESTs:  
(10142..9722) AA470903 ne20h12.r1 NCI CGAP C03 Homo  
sapiens cDNA clone IMAGE:881831. Score: 807 Identity:  
414/422 (98%).  
misc\_feature  
10147..9727) AA642553 nq7a11.r1 NCI CGAP Pr22 Homo  
sapiens cDNA clone IMAGE:1157948. Score: 817 Identity:  
416/420 (99%).- (10218..9777) AA125808 z129e12.r1 Soares  
pregnant uterus NBHPU Homo sapiens cDNA clone 503374 3'.  
Score: 851 Identity: 435/441 (98%).- (10219..9803) AA483290  
ne39b09.r1 NCI CGAP C03 Homo sapiens cDNA clone  
IMAGE:899705 similar to TR:G191151 G191151 PRO-ALPHA-1  
TYPE V COLLAGEN. Score: 804 Identity: 412/415  
(99%).- Other EST matches:  
AA909760"

misc\_feature 10481..11538  
/note="PDS similarity to overlapping ESTs:  
Query Match 1.7%; Score 52; DB 8; Length 42572;  
Best Local Similarity 100.0%; Pred. No. 5.5e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 3071 CAAGATTGTGCACCTGCACCTCGGCAACAGACCAAGACTCTCTTC 3122  
Db 31346 CAAGATTGTGCACCTGCACCTCGGCAACAGACCAAGACTCTCTTC 31397  
RESULT 66  
AC016627  
LOCUS AC016627 44449 bp DNA linear PRI 29-SEP-2000  
DEFINITION Homo sapiens chromosome 19 clone LTNLR-251F1, complete sequence.  
AC016627  
VERSION AC016627.6 GI:10337634  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominae; Homo.  
REFERENCE 1 (bases 1 to 44449)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
Direct Submission  
Unpublished  
2 (bases 1 to 44449)  
REFERENCE  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (04-DEC-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 44449)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
Direct Submission  
Submitted (29-SEP-2000) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
On Sep 29, 2000 this sequence version replaced gi:7711581.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.sbgc.stanford.edu  
Quality: Phrap Quality >=40 99.6% of Sequence;  
Estimated Total Number of Errors is 0.3.  
FEATURES  
source  
1..44449  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clone="LTNLR-251F1"  
ORIGIN  
Query Match 1.7%; Score 52; DB 8; Length 44449;  
Best Local Similarity 100.0%; Pred. No. 5.5e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 3071 CAAGATTGTGCACCTGCACCTCGGCAACAGACCAAGACTCTCTTC 3122  
Db 9282 CAAGATTGTGCACCTGCACCTCGGCAACAGACCAAGACTCTCTTC 9333  
RESULT 67  
CR936927  
LOCUS CR936927 44783 bp DNA linear PRI 03-MAR-2005  
DEFINITION Human DNA sequence from clone DAMC-32J6 on chromosome 6, complete  
sequence.  
ACCESSION CR936927  
VERSION CR936927.1 GI:60495271  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens



```

/sub clone="UMGC:Y55c068"
/cell_line="CGM1"
/clone_lib="Wash U YAC library"
repeat_region
/rpt_family="Alu"
143. .751
repeat_region
/rpt_family="Alu"
1012. .1311
repeat_region
/rpt_family="Alu"
1720. .2002
repeat_region
/rpt_family="Alu"
complement(2189. .2306)
/rpt_family="LTR8"
2390. .2642
repeat_region
/rpt_family="Alu"
2823. .3142
repeat_region
/rpt_family="Alu"
complement(3342. .3488)
/rpt_family="LTR8"
repeat_region
complement(3794. .4088)
/rpt_family="Alu"
complement(4597. .4976)
/rpt_family="Alu"
repeat_region
complement(5537. .5816)
/rpt_family="Alu"
repeat_region
complement(6041. .6333)
/rpt_family="Alu"
complement(6479. .12641)
/rpt_family="LTR"
repeat_region
complement(11144. .11222)
/rpt_family="MER25"
repeat_region
complement(11442. .11610)
/rpt_family="MER25"
1344T. .13727
repeat_region
/rpt_family="Alu"
14018. .14195
repeat_region
/rpt_family="Alu"
complement(14307. .15266)
/rpt_family="LTR12"
15725. .16027
repeat_region
/rpt_family="Alu"
16086. .16350
repeat_region
/rpt_family="Alu"
16386. .16701
repeat_region
/rpt_family="Alu"
17427. .17724
repeat_region
/rpt_family="Alu"
complement(18377. .19150)
/rpt_family="MER4"
repeat_region
complement(18685. .18976)
/rpt_family="Alu"
repeat_region
complement(19202. .19442)
/rpt_family="Alu"
19789. .19923
repeat_region
/rpt_family="Alu"
complement(20024. .20180)
/rpt_family="Alu"
repeat_region
complement(20422. .20679)
/rpt_family="Alu"
repeat_region
complement(20804. .21089)
/rpt_family="Alu"
repeat_region
complement(21281. .21568)
/rpt_family="Alu"
repeat_region
complement(21913. .22032)
/rpt_family="Alu"
repeat_region
complement(22541. .22790)
/rpt_family="MER42"
repeat_region
complement(22816. .22968)
/rpt_family="Alu"
23481. .23767
repeat_region
/rpt_family="Alu"
complement(25929. .26198)
/rpt_family="Alu"
26298. .27083
repeat_region
/rpt_family="Alu"

```

```

repeat_region complement(27761. .28381)
/rpt_family="Alu"
repeat_region 28631. .28906
/rpt_family="Alu"
repeat_region complement(29405. .29699)
/rpt_family="Alu"
repeat_region 29951. .30236
/rpt_family="Alu"
repeat_region 31046. .31644
/rpt_family="Alu"
repeat_region 33178. .33466
/rpt_family="Alu"
repeat_region 33713. .34001
/rpt_family="Alu"
repeat_region complement(34116. .34209)
/rpt_family="Alu"
repeat_region complement(34426. .34369)
/rpt_family="Alu"
repeat_region 34841. .35164
/rpt_family="LTR5"
repeat_region 35170. .35277
/rpt_family="LTR5"
repeat_region complement(36065. .36319)
/rpt_family="Alu"
repeat_region 37201. .37670
/rpt_family="MER1"
repeat_region complement(37323. .37418)
/rpt_family="MER1"
repeat_region 38132. .38421
/rpt_family="Alu"
repeat_region complement(39377. .39729)
/rpt_family="MSTAR"
repeat_region complement(40127. .40416)
/rpt_family="Alu"
repeat_region 41331. .41643
/rpt_family="Alu"
repeat_region 44180. .44436
/rpt_family="Alu"
repeat_region complement(44598. .44804)
/rpt_family="Alu"
repeat_region 44890. .45211
/rpt_family="Alu"

```

```

Query Match 1.7%; Score 52; DB 8; Length 46201;
Best Local Similarity 100.0%; Pred. No. 5,5e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3071 CAAGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACCAAGACTCTGTCTC 3122
DB 20873 CAAGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACCAAGACTCTGTCTC 20822

```

```

RESULT 69
AC108747/c 46509 bp DNA linear PRI 21-JUN-2002
LOCUS Homo sapiens 3 BAC RP11-268B23 (Roswell Park Cancer Institute Human
DEFINITION BAC library) complete sequence.
ACCESSION AC108747.5 GI:21535890
VERSION AC108747.5
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 46509)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Albrooke,S.L., Amaralunge,H.C., Are,J.R., Ayale,M., Banks,T.,
Barbata,J., Benton,J., Blimke,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,U., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,

```





```

repeat_region      18116..18138
                    /rpt_family="AT-rich"
repeat_region      18713..18768
                    /rpt_family="(TATATG)n"
repeat_region      18770..18835
                    /rpt_family="GA-rich"
repeat_region      19136..19424
                    /rpt_family="AluSc"
repeat_region      complement(20236..20362)
                    /rpt_family="MIR"
repeat_region      21144..21183
                    /rpt_family="(CA)n"
repeat_region      21974..22010
                    /rpt_family="(CA)n"
repeat_region      22312..22559
                    /rpt_family="(TG)n"
repeat_region      complement(22584..22635)
                    /rpt_family="MER31B"
repeat_region      24689..24727
                    /rpt_family="MER102"
repeat_region      24728..25024
                    /rpt_family="AluSc"
repeat_region      25025..25240
                    /rpt_family="MER102"
STS
25577..25763
                    /standard_name="60878"
STS
26576..26677
                    /standard_name="62312"
repeat_region      26649..26742
                    /rpt_family="MIR"
repeat_region      27097..27117
                    /rpt_family="AT-rich"
repeat_region      complement(27489..27606)
                    /rpt_family="MER5A"
repeat_region      27758..27935

```

Query Match 1.7%; Score 52; DB 8; Length 46509;

Best Local Similarity 100.0%; Pred. No. 5.5e-16; Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      3071 CAAGATTGTGCACCTGCAGCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122
         |||||||
Db       45643 CAAGATTGTGCACCTGCAGCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 45592

```

```

RESULT 70
LOCUS    CR759912          47462 bp      DNA      linear      PRI 16-OCT-2004
DEFINITION Human DNA sequence from clone DAAP-394023 on chromosome 6, complete
sequence.
ACCESSION CR759912
VERSION   CR759912.6   GI:54262068
KEYWORDS  HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
           Hominiidae; Homo.
REFERENCE 1 (bases 1 to 47462)
AUTHORS   Gray,E.
TITLE     Direct Submission
JOURNAL   Submitted (16-OCT-2004) Wellcome Trust Sanger Institute, Hinxton,
           Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
           humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
           On Oct 16, 2004 this sequence version replaced gi:154021829.
COMMENT   ----- Genome Center
           Center: Wellcome Trust Sanger Institute
           Center code: SC
           Web site: http://www.sanger.ac.uk
           Contact: humquery@sanger.ac.uk
           -----
           During sequence assembly data is compared from overlapping clones.
           Where differences are found these are annotated as variations

```

together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: EMBL, SW, SWISSPROT, Tr, TREMBL, WP, WORMPEP, Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs constructed by the MGC HaploTYPE Consortium and collaborators. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6/MHC> DAAP-394023 is from a Apd human bac - B Lymphoblastoid Cell Line library VECTOR: pBelOBAC11.

#### FEATURES

```

source
1..47462
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="6"
   /clone="DAAP-394023"
   /clone_11b="APD"

```

#### ORIGIN

Query Match 1.7%; Score 52; DB 8; Length 47462;

Best Local Similarity 100.0%; Pred. No. 5.5e-16; Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY      3071 CAAGATTGTGCACCTGCAGCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122
         |||||||
Db       5661 CAAGATTGTGCACCTGCAGCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 5712

```

```

RESULT 71
LOCUS    AL390837          51090 bp      DNA      linear      HTG 10-JUN-2001
DEFINITION Homo sapiens chromosome 1 clone RP5-11168A16 map p36.2, 5 unordered
pieces.
ACCESSION AL390837
VERSION   AL390837.7   GI:13785281
KEYWORDS  HTG; HTGS PHASE1; HTGS_CANCELLED.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
           Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
           Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
           Hominiidae; Homo.
REFERENCE 1
AUTHORS   Plumb,B.
TITLE     Direct Submission
JOURNAL   Submitted (09-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire,
           CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
           requests: clonerequests@sanger.ac.uk
           On Apr 24, 2001 this sequence version replaced gi:13750888.
COMMENT   ----- Genome Center
           Center: Sanger Centre
           Center code: SC
           Web site: http://www.sanger.ac.uk
           Contact: humquery@sanger.ac.uk
           ----- Project Information
           Center project name: d11168A16
           ----- Summary Statistics
           Assembly program: XGAP; version 4.5
           Sequencing vector: plasmid; L08752; 100% of reads
           Chemistry: Dye-terminator Big Dye; 100% of reads

```

Consensus quality: 47312 bases at least Q40  
 Consensus quality: 47821 bases at least Q30  
 Consensus quality: 48428 bases at least Q20  
 Insert size: 50630; sum-of-contigs  
 Insert size: 85481; 31.1% error; agarose-fp  
 Quality coverage: 6.45x in Q20 bases; sum-of-contigs Quality  
 coverage: 4.61x in Q20 bases; agarose-fp

NOTE: This is a 'working draft' sequence. It currently  
 consists of 5 contigs. The true order of the pieces  
 is not known and their order in this sequence record is  
 arbitrary. Gaps between the contigs are represented as  
 runs of N, but the exact sizes of the gaps are unknown.  
 This record will be updated with the finished sequence  
 as soon as it is available and the accession number will  
 be preserved.

1 34808: contig of 34808 bp in length  
 \* 34809: gap of 100 bp  
 \* 34909: contig of 2244 bp in length  
 \* 37153: gap of 100 bp  
 \* 37253: contig of 2473 bp in length  
 \* 39726: gap of 100 bp  
 \* 39826: contig of 5421 bp in length  
 \* 45247: gap of 100 bp  
 \* 45347: 51090: contig of 5744 bp in length.

## FEATURES

source  
 1. 51090  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="1"  
 /map="p36.2"  
 /clone="RP5-1168A16"  
 /clone\_11b="RP1-5"  
 1. 34808  
 /note="assembly\_fragment:00349"  
 34909. 37152  
 /note="assembly\_fragment:00417"  
 37253. 39725  
 /note="assembly\_fragment:01180"  
 39826. 45246  
 /note="assembly\_fragment:00548"  
 fragment chain:1  
 clone end:17  
 vector\_side:right"  
 45347. 51090  
 /note="assembly\_fragment:00789"  
 fragment chain:1  
 clone end:SP6  
 vector\_side:right"

## ORIGIN

Query Match 1.7%; Score 52; DB 14; Length 51090;  
 Best Local Similarity 100.0%; Pred. No. 5.5e-16;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAAGTGTGCACTGCACTGCGCAACAGAGCAAGACTGTGTCTC 3122  
 Db 32300 CAAAGTGTGCACTGCACTGCGCAACAGAGCAAGACTGTGTCTC 32351

RESULT 72  
 ACT103989 59744 bp DNA linear HTG 01-DEC-2001  
 LOCUS Homo sapiens chromosome 18 clone RP11-407C18 map 18, LOW-PASS  
 DEFINITION  
 SEQUENCE SAMPLING.  
 AC103989  
 AC103989.1 GT:17223277  
 VERSION  
 KEYWORDS HTG; HTGS\_PHASE0.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

## REFERENCE

1 (bases 1 to 59744)  
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
 Homo sapiens chromosome 18, clone RP11-407C18  
 Unpublished  
 2 (bases 1 to 59744)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,  
 Anderson, S., Barron, N., Bastien, V., Boguslavsky, L., Boukhaltier, B.,  
 Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B.,  
 Choquet, Y., Colangelo, M., Collins, S., Collinmore, A., Cook, A.,  
 Cooke, P., DeArillano, K., Dewar, K., Diaz, J.S., Dodge, S., Fero, S.,  
 Ferreira, P., Fitzhugh, W., Gage, D., Gallejan, J., Gardyna, S.,  
 Glade, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
 Hagos, B., Harford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
 Jones, C., Kanat, A., Karatas, A., Kelle, C., Larocque, K.,  
 Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,  
 Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C.,  
 McCarthy, M., McEwan, P., McKernan, K., McPeck, R., Meldrum, J.,  
 Menais, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C.,  
 Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D.,  
 Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V.,  
 Raymond, C., Rector, R., Rieback, M., Riley, R., Rise, C., Rogov, P.,  
 Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R.,  
 Seaman, S., Severy, P., Spencer, B., Stange-Thoman, N., Stojanovic, N.,  
 Strauss, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,  
 Topham, K., Travers, M., Travis, N., Trigilio, J., Vasilev, H.,  
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G.,  
 Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.

## TITLE

Direct Submission  
 Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 All repeats were identified using RepeatMasker:  
 Smit, A.P.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RW/RepeatMasker.html

## COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: 121885  
 Center clone name: 407\_C\_18

NOTE: This record contains 75 individual  
 \* sequencing reads that have not been assembled into  
 \* contigs. Runs of N are used to separate the reads  
 \* and the order in which they appear is completely  
 \* arbitrary. Low-pass sequence sampling is useful for  
 \* identifying clones that may be gene-rich and allows  
 \* overlap relationships among clones to be deduced.  
 \* However, it should not be assumed that this clone  
 \* will be sequenced to completion. In the event that  
 \* the record is updated, the accession number will  
 \* be preserved.

1 677: contig of 677 bp in length  
 \* 678: gap of 100 bp  
 \* 778: 1448: contig of 671 bp in length  
 \* 1449: gap of 100 bp  
 \* 1549: 2244: contig of 636 bp in length  
 \* 2245: gap of 100 bp  
 \* 2345: 3032: contig of 688 bp in length  
 \* 3033: gap of 100 bp  
 \* 3133: 3820: contig of 688 bp in length  
 \* 3821: gap of 100 bp  
 \* 3920: gap of 100 bp  
 \* 3921: 4705: contig of 685 bp in length  
 \* 4706: gap of 100 bp  
 \* 5411: contig of 706 bp in length  
 \* 5412: gap of 100 bp  
 \* 5512: 6210: contig of 639 bp in length  
 \* 6211: gap of 100 bp  
 \* 6311: 7016: contig of 706 bp in length  
 \* 7017: gap of 100 bp  
 \* 7117: 7837: contig of 721 bp in length

```

* 7838 7937: gap of 100 bp
* 7938 8620: contig of 683 bp in length
* 8621 8720: gap of 100 bp
* 8721 9421: contig of 701 bp in length
* 9422 9521: gap of 100 bp
* 9522 10227: contig of 706 bp in length
* 10228 10327: gap of 100 bp
* 10328 11026: contig of 699 bp in length
* 11027 11126: gap of 100 bp
* 11127 11821: contig of 695 bp in length
* 11822 11921: gap of 100 bp
* 11922 12614: contig of 693 bp in length
* 12615 12714: gap of 100 bp
* 12715 13404: contig of 690 bp in length
* 13405 13504: gap of 100 bp
* 13505 14209: contig of 705 bp in length
* 14210 14309: gap of 100 bp
* 14310 14988: contig of 679 bp in length
* 14989 15088: gap of 100 bp
* 15089 15793: contig of 705 bp in length
* 15794 15893: gap of 100 bp
* 15894 16524: contig of 631 bp in length
* 16525 17326: contig of 702 bp in length
* 17327 17426: gap of 100 bp
* 17427 18118: contig of 692 bp in length
* 18119 18218: gap of 100 bp
* 18219 18917: contig of 699 bp in length
* 18918 19017: gap of 100 bp
* 19018 19711: contig of 694 bp in length
* 19712 19811: gap of 100 bp
* 19812 20510: contig of 699 bp in length
* 20511 20610: gap of 100 bp
* 20611 21289: contig of 679 bp in length
* 21290 21389: gap of 100 bp
* 21390 22091: contig of 702 bp in length
* 22092 22191: gap of 100 bp
* 22192 22899: contig of 708 bp in length
* 22900 22999: gap of 100 bp
* 23000 23695: contig of 696 bp in length
* 23696 23795: gap of 100 bp
* 23796 24494: contig of 699 bp in length
* 24495 24594: gap of 100 bp
* 24595 25287: contig of 693 bp in length
* 25288 25387: gap of 100 bp
* 25388 26065: contig of 678 bp in length
* 26066 26165: gap of 100 bp
* 26166 26868: contig of 703 bp in length
* 26869 26968: gap of 100 bp
* 26969 27734: contig of 766 bp in length
* 27735 27834: gap of 100 bp
* 27835 28629: contig of 795 bp in length
* 28630 28729: gap of 100 bp
* 28730 29432: contig of 703 bp in length
* 29433 29532: gap of 100 bp
* 29533 30211: contig of 679 bp in length
* 30212 30311: gap of 100 bp
* 30312 31011: contig of 700 bp in length
* 31012 31111: gap of 100 bp
* 31112 31800: contig of 689 bp in length
* 31801 31900: gap of 100 bp
* 31901 32595: contig of 695 bp in length
* 32596 32695: gap of 100 bp
* 32696 33386: contig of 691 bp in length
* 33387 33486: gap of 100 bp
* 33487 34177: contig of 691 bp in length
* 34178 34277: gap of 100 bp
* 34278 34944: contig of 667 bp in length
* 34945 35044: gap of 100 bp
* 35045 35745: contig of 701 bp in length
* 35746 35845: gap of 100 bp
* 35846 36549: contig of 704 bp in length
* 36550 36649: gap of 100 bp

```

```

* 36550 37364: contig of 715 bp in length
* 37365 37464: gap of 100 bp
* 37465 38174: contig of 710 bp in length
* 38175 38274: gap of 100 bp
* 38275 38976: contig of 702 bp in length
* 38977 39076: gap of 100 bp
* 39077 39778: contig of 702 bp in length
* 39779 39878: gap of 100 bp
* 39879 40564: contig of 686 bp in length
* 40565 40664: gap of 100 bp
* 40665 41354: contig of 690 bp in length
* 41355 41454: gap of 100 bp
* 41455 42144: contig of 690 bp in length
* 42145 42244: gap of 100 bp
* 42245 42951: contig of 707 bp in length
* 42952 43051: gap of 100 bp
* 43052 43759: contig of 708 bp in length
* 43760 43859: gap of 100 bp
* 43860 44560: contig of 701 bp in length
* 44561 44560: gap of 100 bp
* 44561 45366: contig of 706 bp in length
* 45367 45466: gap of 100 bp
* 45467 46142: contig of 676 bp in length
* 46143 46242: gap of 100 bp
* 46243 46952: contig of 710 bp in length
* 46953 47052: gap of 100 bp
* 47053 47750: contig of 698 bp in length
* 47751 47850: gap of 100 bp
* 47851 48546: contig of 696 bp in length
* 48547 48646: gap of 100 bp
* 48647 49342: contig of 696 bp in length
* 49343 49442: gap of 100 bp
* 49443 50120: contig of 678 bp in length
* 50121 50220: gap of 100 bp
* 50221 50926: contig of 706 bp in length
* 50927 51026: gap of 100 bp
* 51027 51714: contig of 688 bp in length
* 51715 51814: gap of 100 bp
* 51815 52518: contig of 704 bp in length
* 52519 52618: gap of 100 bp
* 52619 53329: contig of 711 bp in length
* 53330 53429: gap of 100 bp
* 53430 54114: contig of 685 bp in length
* 54115 54214: gap of 100 bp

```

Query Match 1.7%; Score 52; DB 14; Length 59744;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-16;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGGCAGCAGCTCCAGGCTGGGCAAGACAGACAGACTCTGCTC 3122  
 Db 38530 CAAGATTGTGGCAGCTCCAGGCTGGGCAAGACAGACAGACTCTGCTC 38581

RESULT 73  
 LOCUS AL499606/c 62761 bp DNA linear PRI 18-MAY-2005  
 DEFINITION Human DNA sequence from clone RP11-13016 on chromosome 6 contains  
 two CpG islands, complete sequence.

ACCESSION AL499606  
 VERSION AL499606.18 GI:17973944  
 KEYWORDS HMG, Cpg island, HFM1, winged helix/forhead transcription factor.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 62761)  
 Corby, N.  
 Direct Submission  
 Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,  
 Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegase@sanger.ac.uk

TITLE  
 JOURNAL



```

source          1..64056
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="19"
                /clone="CTD-2094A2"

ORIGIN
Query Match      1.7%; Score 52; DB 8; Length 64056;
Best Local Similarity 100.0%; Pred. No. 5.6e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCACCTGCACCTCCAGCCTGGGCAACAGACAGACACTCTGTCTC 3122
DB 38283 CAAGATTGTGCACCTGCACCTCCAGCCTGGGCAACAGACAGACACTCTGTCTC 38334

RESULT 75
AF411057/c      66887 bp   DNA      linear   PRI 13-DEC-2001
LOCUS           Homo sapiens clone BAC.22606 CD28 antigen (CD28) gene, exons 1, 2
DEFINITION      AF411057
ACCESSION       AF411057
VERSION         AF411057.1 GI:17646225
KEYWORDS
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
                Hominoidea; Homo.
REFERENCE
AUTHORS         Ling,V., Wu,P.W., Finnerty,H.F., Agostino,M.J., Graham,J.R.,
                Chen,S., Juszeff,J.M., Fish,G.J., Miller,C.P. and Collins,M.
                Assembling and Annotation of Human Chromosome 2q33 Sequence
                Containing the CD28, CTAA4, and ICOS Gene Cluster: Analysis by
                Computational, Comparative, and Microarray Approaches
                Genomics 78 (3), 155-168 (2001)
TITLE           11735222
                2 (bases 1 to 66887)
JOURNAL         Ling,V., Wu,P.W., Finnerty,H.F., Agostino,M.J., Graham,J.R.,
                Chen,S., Juszeff,J., Fish,G.J., Miller,C.P. and Collins,M.
                Direct Submission
                Submitted (15-AUG-2001) Department of Immunology, Genetics
                Institute/Myech Research, 200 Cambridge Park Drive, Cambridge, MA
                02140, USA
FEATURES
source          Location/Qualifiers
                1..66887
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="2"
                /map="2q33"
                /clone="BAC.22606"
                7838..8329
                /note="similar to NADH:ubiquinone oxidoreductase"
                42348..>66887
                /gene="CD28"
                /join(42348..42621,62505..62861,65540..>65664)
                /gene="CD28"
                /product="CD28 antigen"
                42348..42621
                /gene="CD28"
                /number=1
                42348..42569
                /gene="CD28"
                /join(42570..42621,62505..62861,65540..>65664)
                /gene="CD28"
                /codon_start=1
                /product="CD28 antigen"
                /protein_id="AA140931.1"
                /db_xref="GI:17646226"
                /translation="MRLRLAALNLPSPISIVTGNKILVKQSPMLVAYDNAVNLSCKYSY
                NLFRRFRASLHKGLDSAVEVCVYGVNYSQQLQVYSKTFNCQKLGNSVTFLYQNL

```

```

exon
                YVNOTDIYFCKIEVMYDPPLNDNEKNGTIIHVXGKHLCPSPLPFGSPKPFVWLWVG
                GVLACXSLVTVAFIIFM"
                62505..62861
                /gene="CD28"
                /number=2
                65540..65664
                /gene="CD28"
                /number=3

exon

ORIGIN
Query Match      1.7%; Score 52; DB 8; Length 66887;
Best Local Similarity 100.0%; Pred. No. 5.6e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCACCTGCACCTCCAGCCTGGGCAACAGACAGACACTCTGTCTC 3122
DB 13733 CAAGATTGTGCACCTGCACCTCCAGCCTGGGCAACAGACAGACACTCTGTCTC 13682

RESULT 76
HSJ735P11/c     68959 bp   DNA      linear   PRI 18-MAY-2005
LOCUS           Human DNA sequence from clone RP4-735P11 on chromosome Xp11.3-11.4
DEFINITION      HSJ735P11
ACCESSION       AL049735
VERSION         AL049735.4 GI:5596978
KEYWORDS        HTG.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
                Hominoidea; Homo.
REFERENCE
AUTHORS         Bird,C.
                Direct Submission
                Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
                Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
                On Jul 27, 1999 this sequence version replaced gi:5101820.
                The following abbreviations are used to associate primary accession
                numbers given in the feature table with their source databases:
                Emi.: EMBL; Sw.: SWISSPROT; Tr.: TREMBL; Wp.: WORMPEP; Information
                on the WORMPEP database can be found at
                http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
                was generated from part of bacterial clone contigs of human
                chromosome X, constructed by the Sanger Centre chromosome X Mapping
                Group. Further information can be found at
                http://www.sanger.ac.uk/HGP/ChrX
                RP4-735P11 is from the library RCI-4 constructed by the group of
                Piter de Jong. For further details see
                http://www.chori.org/bacpac/home.htm
                VECTOR: pCYPAC2
                ----- Genome Center
                Center: Wellcome Trust Sanger Institute
                Center code: SC
                Web site: http://www.sanger.ac.uk
                Contact: vegas@sanger.ac.uk
                -----
                This sequence was finished as follows unless otherwise noted: all
                regions were either double-stranded or sequenced with an alternate
                chemistry or covered by high quality data (i.e., phred quality >=
                30); an attempt was made to resolve all sequencing problems, such
                as compressions and repeats; all regions were covered by at least
                one subclone; and the assembly was confirmed by restriction digest,
                except on the rare occasion of the clone being a YAC.
                Location/Qualifiers
                1..68959
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="RZPD:RCP1704P11735"
                /db_xref="taxon:9606"
                /map="p11.3-11.4"
                /clone="RP4-735P11"

```

ORIGIN /clone\_11b="RPCT-4"

Query Match 1.74; Score 52; DB 8; Length 68959;  
Best Local Similarity 100.0%; Pred. No. 5,6e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2889 GAGGACAGTGTGATCAGGAGGACAGAGTTCAGACGAGCTGCGCAACAT 2940  
18500 GAGGACAGTGTGATCAGGAGGACAGAGTTCAGACGAGCTGCGCAACAT 18449

RESULT 77  
AC073488/c

LOCUS Homo sapiens X BAC RP11-126022 (Rosewell Park Cancer Institute Human  
DEFINITION BAC library) complete sequence.  
AC073488  
AC073488.23 GI:27436742  
HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.

REFERENCE  
AUTHORS

1 (bases 1 to 70393)  
Muzny, D.M., Adame, C., Adio-Oduola, B., Ali-ouman, F.R., Allen, C.,  
Albrooks, S.L., Amaratunga, H.C., Are, J.R., Banks, T., Barbier, J.,  
Benton, J., Bimagne, K., Blankenburg, K., Bonnin, D., Bouck, J.,  
Bowle, S., Brileva, M., Brown, B., Brown, M., Bryant, N.P., Buhay, C.,  
Burch, P., Burkett, C., Burrell, K.L., Byrd, N.C., Carron, T.F.,  
Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, R.,  
Chen, Z., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C.,  
Coyle, M.D., Dethorne, S.R., David, R., Davila, M.L., Davis, C.,  
Davy, Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, O.,  
Dem, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Drexler, H.,  
Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C.,  
Elhaj, C., Escotto, M., Falls, T., Ferraguto, D., Flagg, N., Ford, J.,  
Foster, P., Frantz, P., Gabisi, A., Gao, J., Garcia, A., Garner, T.,  
Garrar, N., Gill, R., Gorrell, J.H., Guevara, W., Gunaratne, P., Hale, S.,  
Hamilton, K., Harris, C., Harris, K., Hart, M., Havlak, P., Hawes, A.,  
Hernandez, J., Hernandez, O., Hodgson, A., Hognes, M., Holloway, C.,  
Hollins, B., Homsl, F., Howard, S., Huber, J., Hulyk, S., Hume, J.,  
Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolivet, S.,  
Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J.,  
Kovar, C., Kratovic, J., Kureshi, A., Landry, N., Leal, B., Lewis, L.C.,  
Lewis, L., Li, J., Li, Z., Lichtarge, O., Lieu, C., Liu, J., Liu, W.,  
Luisseg, H., Lozano, R.J., Lu, X., Lucier, A., Lucier, R., Luna, R.,  
Ma, J., Maheshwari, M., Mapua, P., Martin, R., Martindale, A.,  
Martinez, E., Massey, E., Mawhinney, E., McLeod, M.P., Meador, M.,  
Mei, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,  
Morjan, M., Morris, S., Moser, M., Neal, D., Newton, J., Newton, N.,  
Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nwokenkwo, S.,  
Ogulu, M., Okunoye, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B.,  
Peery, J., Peery, L., Peters, L., Pickens, R., Primus, E., Pu, L.L.,  
Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojupokan, I., Rolfe, M.,  
Ruiz, S., Savery, G., Scherer, S., Scott, G., Shen, H., Shoochari, N.,  
Sisson, I., Sodergren, B., Sonalke, T., Sparks, A., Stanley, R.,  
Stone, H., Sutton, A., Svatek, A., Taber, P., Tamerisa, A., Tamerisa, K.,  
Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,  
Thomas, S., Umanai, K., Vasquez, L., Vera, V., Villalob, D., Vinson, R.,  
Wang, O., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,  
Watling, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,  
Weinstock, G. and Gibbs, R.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

3 (bases 1 to 70393)  
Morley, K.C.  
Direct Submission  
Submitted (29-JUN-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
4 (bases 1 to 70393)  
Morley, K.C.  
Direct Submission  
Submitted (31-DEC-2002) Human Genome Sequencing Center, Department  
of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Dec 31, 2002 this sequence version replaced gi:21629140.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
[gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

COMMENT

CLONE LENGTH: This sequence does not necessarily represent the  
entire insert of this clone. Overlapping regions of clones are only  
sequenced and submitted once, so the sequence for the remainder of  
the insert may be found in the record for the adjacent clones.  
Overlapping clones are noted at the beginning and end of the  
Features listing.

#### ANNOTATION OF FEATURES:

STS are identified using ePCR (Genome Res. 7:541-550) searches  
of a local database that includes entries from dbSTS, GDB, and  
local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,  
unpublished) for human and mouse sequences.

Genes and region of sequence similarity are identified by BLAST  
(Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the  
EST and cDNA sequences. Genes demonstrate at least two exons  
flanked by consensus splice sites that maintained sequence  
continuity across the splice junctions. Sequences that are not  
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum  
standard of double strand coverage with a minimum of 2 clones and 2  
reads with no ambiguities or 2 chemistries with a minimum of 2  
clones and 3 reads with no ambiguities. If the sequence quality for  
a region does not meet this standard, it will be indicated in the  
annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality  
standards - estimated error rate less than 1 per 10,000 bases.  
Reports of lowest quality individual bases and measures of base  
quality are listed below. Description of the metrics can be found  
at URL:  
<http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht>  
ml.

#### FEATURES

##### source

1..70393  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="X"  
/clone="RP11-126022"  
repeat\_region  
132..259  
/rpt\_family="7SLRNA"  
repeat\_region  
260..556  
/rpt\_family="AluSg"  
misc\_feature  
463..10771  
/note="sized by restriction digest and formal approved"  
function="unresolved tandem repeat"  
repeat\_region  
558..597  
/rpt\_family="7AGA)n"  
repeat\_region  
598..765  
/rpt\_family="7SLRNA"  
repeat\_region  
766..789  
/rpt\_family="AT-rich"  
repeat\_region  
2233..2309  
/rpt\_family="5S"  
repeat\_region  
2480..2584



```

repeat_region      /rpt_family="SS"
                    4435..4504
repeat_region      /rpt_family="SS"
                    9224..9273
repeat_region      /rpt_family="SS"
                    11202..11248
repeat_region      /rpt_family="SS"
                    12246..12295
repeat_region      /rpt_family="SS"
                    13041..13077
repeat_region      /rpt_family="(TAAAA)n"
                    13850..13876
repeat_region      /rpt_family="AT_rich"
                    14808..14893
repeat_region      /rpt_family="GA-rich"
                    15573..15596
repeat_region      /rpt_family="AT_rich"
                    complement(15634..15894)
repeat_region      /rpt_family="HAL1"
                    complement(15895..16191)
repeat_region      /rpt_family="AluUb"
                    complement(16192..16332)
repeat_region      /rpt_family="HAL1"
                    complement(16511..16685)
repeat_region      /rpt_family="HAL1"
                    complement(16911..17203)
repeat_region      /rpt_family="AluUb"
                    complement(17215..17526)
repeat_region      /rpt_family="AluY"
                    complement(17535..17678)
repeat_region      /rpt_family="Charlies"
                    18285..18430
repeat_region      /rpt_family="FLAM_C"
                    18591..18662
repeat_region      /rpt_family="CT-rich"
                    complement(18777..18925)
repeat_region      /rpt_family="MIR"
                    20368..20668
repeat_region      /rpt_family="AluY"
                    21055..21187
repeat_region      /rpt_family="FLAM_C"
                    21192..21508
repeat_region      /rpt_family="AluUb"
                    complement(21517..21531)
repeat_region      /rpt_family="AluUb"
                    21532..21553
repeat_region      /rpt_family="(TTTA)n"
                    complement(21554..21704)
repeat_region      /rpt_family="AluUb"
                    21705..21727
repeat_region      /rpt_family="(TTTA)n"
                    complement(21728..21847)
repeat_region      /rpt_family="AluUb"
                    complement(21866..22592)
repeat_region      /rpt_family="LMB5"
                    23263..23400
repeat_region      /rpt_family="AluUb"
                    23405..23688
repeat_region      /rpt_family="AluUb"
                    23740..24344
repeat_region      /rpt_family="LMB1"
                    24346..24589
repeat_region      /rpt_family="MLT1D"
                    complement(24590..25032)
repeat_region      /rpt_family="MSTB"
                    25033..25256
repeat_region      /rpt_family="MLT1D"
                    25258..25377

```

```

Query Match      1.7%; Score 52; DB 8; Length 70393;
Best Local Similarity 100.0%; Pred. No. 5.6e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      2889 GAGGACGGTGTGATCATCCAGAGCCAGAGTTGAGACCAAGCGTGGCAACAT 2940
Db      52328 GAGGACGGTGTGATCATCCAGAGCCAGAGTTGAGACCAAGCGTGGCAACAT 52277

RESULT 78
LOCUS   CR388372
DEFINITION Human DNA sequence from clone DADB-87H23 on chromosome 6, complete
sequence.
ACCESSION CR388372
VERSION   CR388372.6
KEYWORDS HTG.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 70448)

REFERENCE
AUTHORS   Wood, J.
TITLE     Direct Submission
JOURNAL   Submitted (17-JUL-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
On Jul 19, 2004 this sequence version replaced gi:49614041.

COMMENT
----- Genome Center
Center: Wellcome Trust Sanger Institute
Genetic code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----

```

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr6> DADB-87H23 is from a DNA-arts DBB human bac library VECTOR: pBel0BAC11.

```

FEATURES
    source
        1..70448
            Location/Qualifiers
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="6"
                /clone="DADB-87H23"
                /clone_lib="DNA-arts-BAC.1-DBB.1"

```

```

ORIGIN
Query Match      1.7%; Score 52; DB 8; Length 70448;
Best Local Similarity 100.0%; Pred. No. 5.6e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy      3071 CAAAGTTGTGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTGTCTC 3122
Db      7753 CAAAGTTGTGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTGTCTC 7804

```





CDS

```

complement (AL049653.7:49897..50102),
complement (AL049653.7:46197..46319),
complement (AL049653.7:45452..45541),
complement (AL049653.7:44490..44582),
complement (AL049653.7:43354..43488),
complement (AL049653.7:42979..43108),
complement (AL049653.7:40722..40972),
complement (AL049653.7:28202..28288),
complement (AL049653.7:25453..25620),
complement (AL049653.7:24850..24962),
complement (AL049653.7:24366..24528),
complement (AL049653.7:22868..22960),
complement (AL049653.7:21366..21512),
complement (AL049653.7:16615..16757),
complement (AL049653.7:11595..11757),
complement (AL049653.7:11312..11457),
complement (53021..53096),complement (42731..42870),
complement (35705..35805),complement (32255..32370),
complement (30547..30624),complement (30227..30344),
complement (25112..25237),complement (24883..25014),
complement (19930..20045),complement (18659..18776),
complement (16108..16356),complement (15317..15417),
complement (14434..14530),complement (14033..14131),
complement (13583..13705),complement (13203..13385),
complement (12589..12723),complement (12201..12375),
complement (10077..10212),complement (7558..7705),
complement (6825..6947),complement (2583..2665),
complement (AL109811.40:111744..111816),
complement (AL109811.40:111161..111235),
complement (AL109811.40:110666..110801),
complement (AL109811.40:109200..109265),
complement (AL109811.40:105997..106077),
complement (AL109811.40:105638..105718),
complement (AL109811.40:104528..104634),
complement (AL109811.40:102883..103848),
/locus_tag="RP4-635E18.1-001"
/feature="PRAP1"
/product="FK506 binding protein 12- rapamycin associated
protein 1"
/notes="match: ESTs; Em:AA262510.1 Em:AA282830.1
Em:AA725390.1 Em:AA737024.1 Em:AA767875.1 Em:AA768596.1
Em:AL044764.1 Em:AL047770.1 Em:AL079658.1 Em:AL120924.1
Em:AL704254.1 Em:AU098639.1 Em:AU135851.1 Em:AW500990.1
Em:AW960244.1 Em:AW976130.1 Em:BE163620.1 Em:BE539676.1
Em:BE560852.1 Em:BE871879.1 Em:BE884193.1 Em:BE898570.1
Em:BE899449.1 Em:BF761404.1 Em:BF761418.1 Em:BF81677.1
Em:BF893441.1 Em:BG250647.1 Em:BG420811.1 Em:BG421771.1
Em:BG768920.1 Em:BG900350.1 Em:BI090093.1 Em:BI58757.1
Em:BI460098.1 Em:BM148769.1 Em:BM149445.1 Em:BM150508.1
Em:BM151339.1 Em:BM153242.1 Em:BM193414.1 Em:BM461878.1
Em:BM470895.1 Em:BM473815.1 Em:BM721256.1 Em:BM724353.1
Em:BM799652.1 Em:BM999881.1 Em:BQ010420.1 Em:BQ024837.1
Em:BQ065337.1 Em:BQ073602.1 Em:BQ040695.1 Em:BQ212803.1 Em:BQ219542.1
Em:BQ221375.1 Em:BQ440695.1 Em:BQ440903.1 Em:BU165455.1
Em:BU431325.1 Em:BU507234.1 Em:BU630556.1 Em:BU741782.1
Em:BU932965.1 Em:BU90957.1 Em:BX437773.1 Em:BX476651.1
Em:BX476627.1 Em:BX481180.1 Em:BX508363.1 Em:CB406561.1
Em:CA429532.1 Em:CA432283.1 Em:CB130997.1 Em:CB426435.1
Em:CD367026.1 Em:CD370823.1 Em:CD540746.1 Em:CD57656.1
Em:CD654263.1 Em:CD657588.1 Em:CD70609.1 Em:T05942.1
match: cDNAs; Em:AF152838.1 Em:J34075.1 Em:T05942.1
join(complement (AL049653.7:71302..71302),
complement (AL049653.7:70539..70647),
complement (AL049653.7:68987..69218),
complement (AL049653.7:68046..68246),
complement (AL049653.7:65893..66027),
complement (AL049653.7:59873..60148),
complement (AL049653.7:59679..59787),
complement (AL049653.7:55168..55354),
complement (AL049653.7:53607..53735),
complement (AL049653.7:52357..52601),
complement (AL049653.7:50456..50671),
complement (AL049653.7:49897..50102),

```

```

Query Match      1.7%; Score 52; DB 8; Length 75525;
Best Local Similarity 100.0%; Pred. No. 5,66-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3071 CAAGTTTGTCACACGCACTCCAGCTGGGCAACAGACCAAGACTGTCTC 3122
Db      7852 CAAGTTTGTCACACGCACTCCAGCTGGGCAACAGACCAAGACTGTCTC 7801

RESULT 81
AC055813
LOCUS      AC055813
DEFINITION Homo sapiens chromosome 17, clone RP11-209W4, complete sequence.
ACCESSION AC055813
VERSION    AC055813.9
KEYWORDS   HTG.

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 80117)
Birren,B., Linton,L., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 17, clone RP11-209W4
Unpublished
2 (bases 1 to 80117)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldoni,J., Barna,N., Bastien,V., Beda,F.,
Boguslavsky,L., Bouhgalter,B., Brown,A., Burkett,G.,
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collimore,A., Cooke,P., Deatrelano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W.,
Galagan,J., Gardyna,S., Glade,S., Goyette,M., Graham,L.,
Grand-Pierre,N., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Iliev,I., Johnson,R., Jones,C., Kam,L., Karatas,A.,
Klein,J., Lakocek,K., Lamazares,R., Landers,T., Lehoczy,J.,
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Margulis,N.,
McCarthy,M., McGowan,P., McGurt,A., McKernan,K., McPheters,R.,
Meldrum,J., Meneus,L., Mihova,T., Miranda,C., Menga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neil,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,K., Talamas,J.,
Teefage-S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,
Vasiliyev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (18-Apr-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 80117)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Allen,N.,
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Bouhgalter,B.,
Brown,A., Canarata,J., Campopiano,A., Chang,J., Chazaro,B.,
Choepel,Y., Colangelo,M., Collins,S., Collimore,A., Cooke,A.,
Cooke,P., Deatrelano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,
Glade,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,
Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,
Jones,C., Kamat,A., Karatas,A., Kelle,C., Lakocek,K.,
Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,
Maclean,C., Macdonald,P., Major,D., Margulis,N., Matthews,C.,

```

McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhthang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnpbach, R., Seaman, S., Severy, P., Sougnuez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travers, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

**TITLE**  
Direct Submission

**JOURNAL**  
Submitted (01-AUG-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

**REFERENCE**  
4 (bases 1 to 80117)

**AUTHORS**  
Birken, B., Linton, L., Nusbbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Barina, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choquel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Johnson, R., Jones, C., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lacroque, K., Lamazeres, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicoli, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhthang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Riback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnpbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travers, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

**TITLE**  
JOURNAL

**COMMENT**  
Submitted (22-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Feb 22, 2002 this sequence version replaced gi:15055348.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L7949  
Center clone name: 209\_M\_4  
-----

Only the last 80.1 Kilobases of this clone are being submitted.  
The remainder overlaps accession number AC004149 [MIGR project L302].

**FEATURES**  
**source**  
1. 80117  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/map="17"  
/clone="RP11-209M4"  
/clone\_lib="RP11-11 Human Male BAC"  
repeat\_region  
complement(146..434)  
/rpt\_family="AluSg"  
repeat\_region  
complement(438..655)  
/rpt\_family="MIR"  
repeat\_region  
complement(974..1129)  
/rpt\_family="MER20"  
repeat\_region  
complement(1270..1438)

/rpt\_family="AluUo"  
1439..1465  
/rpt\_family="TTTTTA)n"  
complement(1466..1584)  
/rpt\_family="AluUo"  
complement(1785..1933)  
/rpt\_family="MER5B"  
2339..2487  
/rpt\_family="MIR"  
complement(2492..2787)  
/rpt\_family="AluSg"  
complement(2796..2902)  
/rpt\_family="MIR"  
complement(2903..3165)  
/rpt\_family="AluSg"  
complement(3166..3193)  
/rpt\_family="MIR"  
3572..3644  
/rpt\_family="MIR"  
4524..4583  
/rpt\_family="MIR3"  
complement(4755..4962)  
/rpt\_family="MIR"  
6397..6738  
/rpt\_family="L2"  
complement(6766..6875)  
/rpt\_family="MLT1H"  
7202..7270  
/rpt\_family="L2"  
7739..7825  
/rpt\_family="L1MC4"  
complement(7826..8122)  
/rpt\_family="AluSc"  
8123..8268  
/rpt\_family="L1MC4"  
8269..8568  
/rpt\_family="AluUo"  
8569..8631  
/rpt\_family="L1MC4"  
complement(9151..9305)  
/rpt\_family="MIR"  
9365..9390  
/rpt\_family="AT rich"  
complement(9405..9708)  
/rpt\_family="AluY"  
complement(9709..9969)  
/rpt\_family="AluUo"  
10197..10458  
/rpt\_family="MLT1K"  
10758..10895  
/rpt\_family="MIR"  
12676..12773  
/rpt\_family="MIR"  
complement(12784..12997)  
/rpt\_family="MIR"  
complement(13220..13513)  
/rpt\_family="AluUo"  
13627..13787  
/rpt\_family="L2"  
complement(13821..13953)  
/rpt\_family="MIR"  
14746..14899  
/rpt\_family="MIR"  
complement(14947..15144)  
/rpt\_family="MIR"  
15520..15668  
/rpt\_family="AluUo"  
15669..15960  
/rpt\_family="AluSx"  
15961..16128  
/rpt\_family="AluUo"  
16208..16400  
/rpt\_family="MIR"

```
repeat_region 16454..16586
repeat_region /rpt_family="I12"
repeat_region 16707..16783

Query Match
Best Local Similarity 1.7%; Score 52; DB 8; Length 80117;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCGAGTGTGATCCTGAGGCGCAGAGTTCAGAGCCCTGGCCCAACAT 2940
DB 54323 GAGGCGAGTGTGATCCTGAGGCGCAGAGTTCAGAGCCCTGGCCCAACAT 54374

RESULT 82
LOCUS BX927220 82724 bp DNA linear PRI 09-APR-2004
DEFINITION Human DNA sequence from clone DAMA-199F12 on chromosome 6, complete
sequence.
ACCESSION BX927220
VERSION BX927220.8 GI:46357732
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 82724)
Wood, J.
Direct Submission
Submitted (09-APR-2004) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
On Apr 9, 2004 this sequence version replaced GI:46016496.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
-----
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL, Sw: SWISSPROT, Tr: TREMBL, Wp: WORMBASE, Information
on the WORMBASE database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormp This sequence
was generated from part of bacterial clone contigs constructed by
the MHC Haplotype Consortium and collaborators. Further information
can be found at
http://www.sanger.ac.uk/HGP/Chr6/MHC
DAMA-199F12 is from the DNA-Arts human BAC library MANN.1 VECTOR:
pBelOBAC11.

FEATURES
source
1..82724
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="6"
/clone="DAMA-199F12"
/clone_lib="DNA-Arts BAC library MANN.1"
```

ORIGIN

```
Query Match
Best Local Similarity 1.7%; Score 52; DB 8; Length 82724;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCACTGCACTCCAGGCTGGGCAACAGAGCAAGCTGTCTC 3122
DB 705 CAAGATTGTGCACTGCACTCCAGGCTGGGCAACAGAGCAAGCTGTCTC 756

RESULT 83
LOCUS AR242855 92139 bp DNA linear PAT 20-DEC-2002
DEFINITION Sequence 1 from patent US 6475739.
ACCESSION AR242855
VERSION AR242855.1 GI:27289517
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.

REFERENCE
1 (bases 1 to 92139)
AUTHORS Brunkow, M.B., Prohl, S., Paepfer, B. and Straehling-Hampton, K.
TITLE Methods for identifying genomic deletions
JOURNAL Patent: US 6475739-A 1 05-NOV-2002;
Celltech R&D, Inc.; Bothell, WA

FEATURES
source
1..92139
/organism="unknown"
/mol_type="genomic DNA"
```

## ORIGIN

```
Query Match
Best Local Similarity 1.7%; Score 52; DB 6; Length 92139;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCGAGTGTGATCCTGAGGCGCAGAGTTCAGAGCCCTGGCCCAACAT 2940
DB 85294 GAGGCGAGTGTGATCCTGAGGCGCAGAGTTCAGAGCCCTGGCCCAACAT 85345
```

```
RESULT 84
LOCUS AX384907 92139 bp DNA linear PAT 19-MAR-2002
DEFINITION Sequence 1 from Patent WO0210455.
ACCESSION AX384907
VERSION AX384907.1 GI:19578035
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE
1
AUTHORS Brunkow, M.B., Prohl, S. and Paepfer, B.
TITLE Methods for identifying genomic deletions
JOURNAL Patent: WO 0210455-A 1 07-FEB-2002;
Celltech R & D, Inc. (US) ; Straehling-Hampton, Karen (US)

FEATURES
source
1..92139
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
```

## ORIGIN

```
Query Match
Best Local Similarity 1.7%; Score 52; DB 6; Length 92139;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCGAGTGTGATCCTGAGGCGCAGAGTTCAGAGCCCTGGCCCAACAT 2940
DB 85294 GAGGCGAGTGTGATCCTGAGGCGCAGAGTTCAGAGCCCTGGCCCAACAT 85345
```



RESULT 85  
LOCUS CR751233/c 92442 bp DNA linear HTG 20-AUG-2004  
DEFINITION Homo sapiens chromosome 6 clone DASS-94P17, WORKING DRAFT SEQUENCE,  
2 unordered pieces.  
ACCESSION CR751233  
VERSION CR751233.1 GI:51491315  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
1 (bases 1 to 92442)  
Sims, S.  
Direct Submission  
Submitted (19-AUG-2004) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:  
humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: humquery@sanger.ac.uk  
----- Project Information  
Center project name: bss94P17  
----- Summary Statistics  
Assembly program: XGAP4; version 4.5  
Chemistry: Dye-terminator; 100% of reads  
Consensus quality: 91963 bases at least Q40  
Consensus quality: 92179 bases at least Q30  
Consensus quality: 92265 bases at least Q20  
Insert size: 92342; sum-of-contigs  
Insert size: 99297; 0.8% error; agarose-fp  
Quality coverage: 6.50x in Q20 bases; sum-of-contigs Quality  
coverage: 6.14x in Q20 bases; agarose-fp  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 2 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 30199: contig of 30199 bp in length  
\* 30200 30299: gap of 100 bp  
\* 30300 92442: contig of 62143 bp in length.  
FEATURES  
source  
1..92442  
/organism="Homo sapiens"  
/mol\_type="Genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone\_1lb="DNA-arcs-BAC.1-SSTO.1"  
/clone\_1lb="DNA-arcs-BAC.1-SSTO.1"  
misc\_feature  
1..30199  
/note="assembly\_fragment:00002.0"  
misc\_feature  
30300..92442  
/note="assembly\_fragment:00375.0"  
ORIGIN  
Query Match 1.7%; Score 52; DB 14; Length 92442;  
Best Local Similarity 100.0%; Pred.No.5.6e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3071 CAAGATTGCGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
|||||  
DB 54359 CAAGATTGCGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 54308  
RESULT 86  
AF397423/c

LOCUS AF397423 93790 bp DNA linear PRI 04-SEP-2002  
DEFINITION Homo sapiens SOST gene, 3' UTR, MEOX1 gene, 5' UTR, and SOST/MEOX1  
intergenic region.  
ACCESSION AF397423  
VERSION AF397423.1 GI:19880618  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
1 (bases 1 to 93790)  
REFERENCE  
AUTHORS Steehling-Hampton,K., Prohl,S., Paepfer,B., Zhao,L., Charney,P.,  
Brown,A., Gardner,J.C., Galas,D., Schatzman,R.C., Beighton,P.,  
Papapoulos,S., Hamersma,H. and Brunkow,M.E.  
A 52-kb deletion in the SOST-MEOX1 intergenic region on 17q12-q21  
is associated with van Buchem disease in the Dutch population  
Am. J. Med. Genet. 110 (2), 144-152 (2002)  
12116252  
2 (bases 1 to 93790)  
REFERENCE  
AUTHORS Steehling-Hampton,K., Prohl,S., Paepfer,B., Zhao,L., Charney,P.,  
Brown,A., Gardner,J.C., Galas,D., Schatzman,R.C., Beighton,P.,  
Papapoulos,S., Hamersma,H. and Brunkow,M.E.  
Direct Submission  
Submitted (06-JUN-2001) Genomics, Celltech R&D Inc., 1631 220th St  
SE, Bothell, WA 98021, USA  
----- Location/Qualifiers  
1..93790  
/organism="Homo sapiens"  
/mol\_type="Genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/map="17q21-q22"  
<1..1612  
/gene="SOST"  
<1..1612  
/gene="SOST"  
1..1612  
/gene="SOST"  
12298..12676  
/standard\_name="D17S951"  
36239..87958  
/note="Van Buchem causative deletion"  
79495..79823  
/standard\_name="D17S1789"  
93762..93790  
/gene="MEOX1"  
93762..93790  
/gene="MEOX1"  
93762..93790  
/gene="MEOX1"  
93762..93790  
/gene="MEOX1"  
93762..93790  
/gene="MEOX1"  
ORIGIN  
Query Match 1.7%; Score 52; DB 8; Length 93790;  
Best Local Similarity 100.0%; Pred.No.5.6e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2889 GAGGAGGTGATCACTGAGGCGAGGAGTTGAGACCAAGCTGGCCAAACAT 2940  
|||||  
DB 8458 GAGGAGGTGATCACTGAGGCGAGGAGTTGAGACCAAGCTGGCCAAACAT 8407  
RESULT 87  
AL645729  
LOCUS AL645729 96217 bp DNA linear PRI 18-MAY-2005  
DEFINITION Human DNA sequence from clone RP11-318E23 on chromosome 1 contains  
the GMEB1 gene for glucocorticoid modulatory element binding  
protein 1, the 5' end of the gene for high-glucose-regulated  
protein 8 (HGR8) and a CpG island, complete sequence.  
ACCESSION AL645729 GI:21322389  
VERSION AL645729.14  
KEYWORDS HTG; GMEB1; HGR8.  
SOURCE Homo sapiens (human)



```

/gene="GMEB1"
/locus_tag="RP11-318E23.1-003"
/standard_name="OTTHUMP00000003781"
/codon_start=1
/product="glucocorticoid modulatory element binding
protein 1"
/protein_id="CAH72427.1"
/db_xref="GI:55666389"
/db_xref="InterPro:IPR010919"
/db_xref="InterPro:IPR000770"
/translation="MANAEVSVPVGVVVVPTGEGNEGNEPDTKTQVILQLOPVQOGL
FIDGHPYRIYAESNNVAVVAVETHTHKIEBGIDTGTIANEDMEIAYITGSES
KAILMKKPGVGINVKYKFNDOILSPGHVLAGSKSTLKMKAIRLKGIMLRKM
DSGQIDPVQHDVCSWTKSTKEDLISARAPVPGQSVVOTPSADGSIOTYAI
ESMERAGLFRNSALTAATVATMTBESYKDSIESEDITMPKGIADVGMEEVCNI
QIEIEBLRGVQRLIAPFVYDAVLNVNVAHTFGIMDTVKKVIDNRNOVEQGEQ
PLVTLTDLEROLEBOKQDRLKSKQTVQNVLMVPSVTPKPKRRLQRPASTVLIS
PSPVQOPFTVISPITTPVQGSFSGMNI PVATLSQSSPVTVHTLPSGQLFRVAT
VVSASSSPDVTYTHSSSLALISSTAMQDSTLGNMTTSPVELVAMEGSLTSAI
QVESTSEDCGTIIEIDPAPDPEADTBEGKAVILTELTEBKVVAEMEHQVHNV
EIVLED"
join(29212..29339,35703..35785,37184..37308,38541..38644,
42526..42663,48037..48168,49791..49928,56119..56241,
59672..60372)
/gene="GMEB1"
/locus_tag="RP11-318E23.1-001"
/standard_name="OTTHUMP00000003779"
/notes="match: proteins: Sw:Q9UL60 Sw:Q9Y692 Tr:Q80Y88"
/codon_start=1
/product="glucocorticoid modulatory element binding
protein 1"
/protein_id="CAH72428.1"
/db_xref="GI:55666390"
/translation="MANAEVSVPVGVVVVPTGEGNEGNEPDTKTQVILQLOPVQOGL
YAESNNVAVVAVETHTHKIEBGIDTGTIANEDMEIAYITGSESXAILMKKPV
DKYINVCYKFNDOILSPGHVLAGSKSTLKMKAIRLKGIMLRKMDSQIDPYQH
CSWTKSTKEDLISARAPVPGQSVVOTPSADGSIOTYAIISSESMEEGLB
NSALTAATVATMTBESYKDSIESEDITMPKGIADVGMEEVCNIQIEIEBLR
VQOQLIAPFVYDAVLNVNVAHTFGIMDTVKKVIDNRNOVEQGEQRLVTLTDLER
QIEBOKQDRLKSKQTVQNVLMVPSVTPKPKRRLQRPASTVLISPSVQOPFT
TVISPITTPVQGSFSGMNI PVATLSQSSPVTVHTLPSGQLFRVATVVSASSSP
DVTYTHSSSLALISSTAMQDSTLGNMTTSPVELVAMEGSLTSAIDAVSTSEDC
GTIIEIDPAPDPEADTBEGKAVILTELTEBKVVAEMEHQVHNVIVLED"
60511
/gene="GMEB1"

polyA_site

Query Match 1.7%; Score 52; DB 8; Length 96217;
- Best Local Similarity 100.0%; Pred. No. 5.6e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTCCTGCACTCCAGCTGGGCAACAGACGACTCTGTCTC 3122
|||||
DB 74409 CAAGATTGTCCTGCACTCCAGCTGGGCAACAGACGACTCTGTCTC 74460
|||||

RESULT 88
AL844213 97630 bp DNA linear PRI 18-MAY-2005
LOCUS AL844213 GI:21953190
DEFINITION Human DNA sequence from clone DAQB-352F19 on chromosome 6 Contains
the TMPDPI pseudogene for thymopoleitin-like 1, the SUC1A2P
pseudogene for succinate-CoA ligase, ADP-forming, beta subunit, the
RANP1 pseudogene for RAN, member RAS oncogene family pseudogene 1,
HLA-B gene for major histocompatibility complex, class I, E, and 6
CPG islands, complete sequence.
ACCESSION AL844213
VERSION AL844213.2 GI:21953190
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
REFERENCE 1 (bases 1 to 97630)

```

```

AUTHORS Tracey A.
TITLE Direct Submission
JOURNAL Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
COMMENT On June 24, 2002 this sequence version replaced gi:21912387.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Bm, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone configs constructed by
the MHC HaploTYPE Consortium and collaborators. Further information
can be found at
http://www.sanger.ac.uk/HGP/Chr6/MHC
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30) ; an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
DAQB-352F19 is from a DNA-arts QBL human bac library VECTOR:
pbeloBAC11.

FEATURES             Location/Qualifiers
     source            1..97630
                     /organism="Homo sapiens"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:9606"
                     /chromosome="6"
                     /clone="DAQB-352F19"
                     /clone_1lb="DNA-arts-BAC.1-QBL.1"
     misc_feature      1
                     /note="Clone left end: DAQB-352F19"
     gene              join(27698..28117,28351..29083)
                     /locus_tag="DAQB-352F19.1-001"
     CDS                join(27698..28117,28351..29083)
                     /locus_tag="DAQB-352F19.1-001"
                     /note="match: proteins: AAC25390 P42167 Q61030 Q61031"
                     /pseudo
                     /codon_start=1
                     /product="thymopoleitin-like 1"
                     /locus_tag="DAQB-352F19.2-001"
                     /pseudo
                     /codon_start=1
                     /product="succinate-CoA ligase, ADP-forming, beta subunit
pseudogene"
                     /locus_tag="DAQB-352F19.3-001"
                     /locus_tag="DAQB-352F19.3-001"
                     /pseudo
                     /locus_tag="DAQB-352F19.3-001"
                     /locus_tag="DAQB-352F19.3-001"
                     /locus_tag="DAQB-352F19.3-001"
                     /note="match: proteins: RANM5923 P51148 Q9F79 Q9VNG6"
                     /pseudo
                     /codon_start=1
                     /product="RAN, member RAS oncogene family pseudogene 1"
                     /note="Clone right end: DAQB-39610"
                     /locus_tag="DAQB-352F19.4-001"
                     /locus_tag="DAQB-352F19.4-001"

```

mRNA  
 join(50669.50755,50886.51155,51400.51675,52297.52572,52697.52813,53564.53596,53701.53743,53909.55108)  
 /gene="HUA-E"  
 /locus\_tag="DAQB-352F19.4-001"  
 /product="major histocompatibility complex, class I, E"  
 /note="match: ESTs: AL545738.1 BM917068.1 BM917589.1 BQ052237.1 BQ689518.1  
 match: CDNA: AK098434.1 BC002578.1 BC004297.1 M20022.1" join(50692.50755,50886.51155,51400.51675,52297.52572,52697.52813,53564.53596,53701.53741)  
 /gene="HUA-E"  
 /locus\_tag="DAQB-352F19.4-001"  
 /standard\_name="OTHTMP0000014549"  
 /note="match: proteins: Q68840 Q31521 Q9B783 Q9MXG3 Q9MX05"  
 /codon\_start=1  
 /product="major histocompatibility complex, class I, E"  
 /protein\_id="CA118596.1"  
 /db\_xref="GI:56206333"  
 /db\_xref="InterPro:IPR001039"  
 /db\_xref="InterPro:IPR003006"  
 /db\_xref="InterPro:IPR007110"  
 /translation="MVDGTLILSLALITQWAGSHSLKTPHTSVSRGRGPRFI SVGYDDTQFVRPDNDASPRMVPRAFMEOGSEYMDRETSARDTAQIFPVNRTL RGYNOSESHSLQMHGCELDGRFLRGYEOFDGKQTLTINEDLRSTAVDTA AQISROKSDASBAEHQRAYLDTCTCYEMAKLLEKKEKTLHLLEPKTVDTHHPISDH EATLQKALGFYPAETITLQQDDGSHGDTDELVERPRGDDGTFQMAVAVPSSGEQ RTCTVQHGELPEPVLIRKMPASOPTPIVGIIAGLVLLGVSAGVAVAIVIRKSS GKGGSYSKAEVSDSNQSGESHSL"  
 55331.55336  
 55354  
 97630  
 /note="Clone\_right\_end: DAQB-352F19"  
 ORIGIN  
 Query Match 1.7%; Score 52; DB 8; Length 97630;  
 Best Local Similarity 100.0%; Pred. No. 5,6e-16;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 3122  
 DB 73737 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 73788  
 RESULT 89  
 AC093203 101769 bp DNA linear HTG 07-MAR-2002  
 LOCUS Homo sapiens chromosome 5 clone CTC-262A8, WORKING DRAFT SEQUENCE,  
 DEFINITION 2 ordered pieces.  
 AC093203 GI:19224825  
 VERSION HTG; PHASE2; HTGS\_DRAFT; HTGS\_ACTIVEFIN.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 101769)  
 REFERENCE AUTHORS DOE Joint Genome Institute.  
 TITLE Sequencing of Human Chromosome 5  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 101769)  
 AUTHORS DOE Joint Genome Institute.  
 TITLE Direct Submission  
 JOURNAL Submitted (16-AUG-2001) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 3 (bases 1 to 101769)  
 REFERENCE 3 (bases 1 to 101769)  
 AUTHORS DOE Joint Genome Institute.  
 TITLE Direct Submission  
 JOURNAL Submitted (07-MAR-2002) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 COMMENT On Mar 7, 2002 this sequence version replaced gi:15426059.

-----Genome Center  
 Center: Joint Genome Institute  
 Center Code: JGI  
 Web site: <http://www.jgi.doe.gov>  
 -----  
 Project Information  
 Center Project Name: 285795  
 Center clone name: CIT-HSPC\_262A8  
 -----  
 Summary Statistics  
 Consensus quality: 101567 bases at least Q40  
 Consensus quality: 101611 bases at least Q30  
 Consensus quality: 101641 bases at least Q20  
 Estimated insert size: 104000; agarose-fp estimation  
 Estimated insert size: 101669; sum-of-contigs estimation  
 Quality coverage: 34.54 in Q20 bases; agarose-fp estimation  
 Quality coverage: 35.33 in Q20 bases; sum-of-contigs estimation.  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 2 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* 1 12059: contig of 12059 bp in length  
 \* 12060 12159: gap of unknown length  
 \* 12160 101769: contig of 89610 bp in length.  
 Location/Qualifiers  
 1..101769  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="5"  
 /clone="CTC-262A8"  
 /clone\_11b="Caltech human BAC library C"  
 12060..12159  
 /estimated\_length=unknown  
 ORIGIN  
 gap  
 Query Match 1.7%; Score 52; DB 14; Length 101769;  
 Best Local Similarity 100.0%; Pred. No. 5,6e-16;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 3122  
 DB 35123 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 35072  
 RESULT 90  
 AC005377 102311 bp DNA linear PRI 03-OCT-2003  
 LOCUS Homo sapiens PAC clone RP5-1136G2 from 7, complete sequence.  
 DEFINITION AC005377  
 ACCESSION AC005377.2 GI:4699956  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 102311)  
 REFERENCE AUTHORS Sulston,J.E. and Wilson,R.  
 TITLE Toward a complete human genome sequence  
 JOURNAL Genome Res. 8 (11), 1097-1108 (1998)  
 PUBMED 9847074  
 REFERENCE 2 (bases 1 to 102311)  
 AUTHORS Bauer,C., Ozanich,A. and Tucci,S.  
 TITLE The sequence of Homo sapiens PAC clone RP5-1136G2  
 JOURNAL Unpublished (2001)  
 REFERENCE 3 (bases 1 to 102311)  
 AUTHORS Waterston,R.H.



```

repeat_region      9946..10330
                    /rpt_family="MALR"
repeat_region      10331..10638
                    /rpt_family="MALR"
repeat_region      10639..10949
                    /rpt_family="Alu"
repeat_region      10950..11092
                    /rpt_family="MALR"
repeat_region      11100..11224
                    /rpt_family="L1"
repeat_region      11381..11577
                    /rpt_family="L1"
repeat_region      11578..11870
                    /rpt_family="Alu"
repeat_region      11871..11950
                    /rpt_family="L1"
repeat_region      12482..12783
                    /rpt_family="Alu"
repeat_region      13738..13897
                    /rpt_family="MERL_type"
repeat_region      13975..14102
                    /rpt_family="MIR"
repeat_region      14199..14263
                    /rpt_family="L2"
repeat_region      14313..14623
                    /rpt_family="Alu"
repeat_region      14698..14798
                    /rpt_family="L1"
repeat_region      14877..15185
                    /rpt_family="Alu"
repeat_region      15190..15283
                    /rpt_family="L1"
repeat_region      15858..16008

```

```

Query Match      1.7%; Score 52; DB 8; Length 10231;
Best Local Similarity 100.0%; Pred. No. 5; 6e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3071 CAAGATTGCGCAGTCACTCCAGCTTGCGCAACAGACAGACTCTGTCTC 3122
Db      75623 CAAGATTGCGCAGTCACTCCAGCTTGCGCAACAGACAGACTCTGTCTC 75572

```

```

RESULT 91
AL359699/c      105902 bp      DNA      linear      PRI 18-MAY-2005
LOCUS      Human DNA sequence from clone RP11-130F5 on chromosome 1 Contains a
DEFINITION      SIAM family member 6 (SIAMF6) pseudogene, the gene for kinase
                  interacting with leukemia-associated gene (KIS) and two CpG
                  islands, complete sequence.
ACCESSION      AL359699 AC068536
VERSION      AL359699.12 GI:18476598
KEYWORDS      HTG; KIS; SIAMF6.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens

```

```

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Homo.
1 (bases 1 to 105902)
REFERENCE
AUTHORS      Chapman,J.
TITLE      Direct Substitution
JOURNAL      Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
                  Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk

```

## COMMENT

On Feb 1, 2002 this sequence version replaced gi:18250758.  
 The following abbreviations are used to associate primary accession  
 numbers given in the feature table with their source databases:  
 Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information  
 on the WORMPEP database can be found at  
[http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence  
 was generated from part of bacterial clone contigs of human  
 chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping  
 Group. Further information can be found at

<http://www.sanger.ac.uk/HGP/Chr1>  
 RP11-130F5 is from the library RP11-11 constructed by the group  
 of Pieter de Jong. For further details see  
<http://www.chori.org/bacpac/home.htm>  
 VECTOR: pBACe3.6  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: <http://www.sanger.ac.uk>  
 Contact: [vegas@sanger.ac.uk](mailto:vegas@sanger.ac.uk)  
 -----  
 Draft Sequence Produced by Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
<http://genome.wustl.edu/gsc/index.shtml>  
 This sequence was finished as follows unless otherwise noted: all  
 regions were either double-stranded or sequenced with an alternate  
 chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
 as compressions and repeats; all regions were covered by at least  
 one subclone; and the assembly was confirmed by restriction digest,  
 except on the rare occasion of the clone being a YAC.  
 Location/Qualifiers  
 1..105902  
   organism="Homo sapiens"  
   mol\_type="Genomic DNA"  
   /db\_xref="taxon:9606"  
   /chromosome="1"  
   /clone="RP11-130F5"  
   /clone\_1fb="RP11-11.1"  
   2000  
   /note="Clone right\_end: RP11-565P22"  
   17865..18222  
   /locus\_tag="RP11-130F5.1-001"  
   /pseudo  
   17885..18222  
   /locus\_tag="RP11-130F5.1-001"  
   /note="match: proteins: Tr:AAQ88526 Tr:Q96DU3 Tr:Q96DVO  
   Tr:Q9ET40"  
   /pseudo  
   /codon\_start=1  
   /product="SIAM family member 6 (SIAMF6) pseudogene"  
   join(70179..70604,72291..72583,73249..73440,76090..76184,  
   84836..84912,85061..85159,90438..90526,94740..101946)  
   /gene="THMK1"  
   /locus\_tag="RP11-130F5.2-001"  
   join(70179..70604,72291..72583,73249..73440,76090..76184,  
   84836..84912,85061..85159,90438..90526,94740..94866)  
   /gene="THMK1"  
   /locus\_tag="RP11-130F5.2-001"  
   /product="U2AF homology motif (UHM) kinase 1"  
   /note="match: cDNAs: Em:ACJ56197.1 Em:AK013347.1  
   Em:BC026046.1 Em:BC058732.1 Em:U70372.2 Em:X58374.1  
   Em:Y10725.2"  
   join(70337..70604,72291..72583,73249..73440,76090..76184,  
   84836..84912,85061..85159,90438..90526,94740..94866)  
   /gene="THMK1"  
   /locus\_tag="RP11-130F5.2-001"  
   /standard\_name="OTTHOMP0000029529"  
   /note="match: proteins: Sw:AAH58732 Sw:P97343 Tr:AA013515  
   Tr:CAD60192 Tr:Q63285 Tr:Q87AS1"  
   /codon\_start=1  
   /product="U2AF homology motif (UHM) kinase 1"  
   /protein\_id="CAH70392.1"  
   /db\_xref="GI:55661991"  
   /db\_xref="GeneW:19683"  
   /db\_xref="GOA:Q87AS1"  
   /db\_xref="InterPro:IPR000504"  
   /db\_xref="InterPro:IPR000719"  
   /db\_xref="InterPro:IPR001245"  
   /db\_xref="InterPro:IPR002290"  
   /db\_xref="InterPro:IPR008271"  
   /db\_xref="InterPro:IPR011009"

```

FEATURES
source
misc_feature
gene
CDS
gene
mRNA
gene
CDS

```



/db\_xref="UniProt/Swiss-Prot:O87AS1"  
/translation="MAGSGCAWCAAPRPLRFGRLMQVSRIGSGSSASVRYVCCG  
NFGSPGALQKFLPRTTGAASAARBYGRKRALBEOQGRNVTLYKVTTFPS  
NPSRCLLELDVSVSELLYSHQGSMNIQICARVLEALFLHEGVADLK  
PNIILMSANECFKLIDFGLSPKQNOVKYIOTDGYRPFELQNCIAQAGLQDTE  
CSAVDMLSGIILLEMFGMKLKTFRSQEWKANSALIDHI PASKAVNAIPAYH  
LMDLKSMLHDDPSRRIPALMALCSPPFSLPPAPPIEDVLMPTVRLNLTDDYL  
EMREEDVDEVEKCKCKYGVSLVLPKENVKPGQVVEAANGDSKAAQKLLTGR  
MDGKRVAVTFPLSAYKRGYKXTLL"  
join(70421..70604,72291..73440,76090..76184,84836..84912,  
85061..85159,90438..90526,94740..96390)  
/gene="UHMK1"  
/locus\_tag="RP11-130F5.2-002"  
join(70421..70604,72291..73440,76090..76184,84836..84912,  
85061..85159,90438..90526,94740..96390)  
/gene="UHMK1"  
/locus\_tag="RP11-130F5.2-002"  
/product="U2AF homology motif (UHMK1) kinase 1"  
/note="match: CDNA8: Em:ALB34136.1"  
96390  
/gene="UHMK1"  
/locus\_tag="RP11-130F5.2-001"  
/note="Clone\_left\_end: RP11-359K18"  
101941..101946  
/gene="UHMK1"  
/locus\_tag="RP11-130F5.2-001"  
101946  
/gene="UHMK1"  
/locus\_tag="RP11-130F5.2-001"  
105902  
/note="Clone\_right\_end: RP11-130F5"  
ORIGIN  
Query Match 1.7%; Score 52; DB 8; Length 105902;  
Best Local Similarity 100.0%; Pred. No. 5,6e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3071 CAAGATTGTGCACTGCACCTCCAGCTGGGCAACAGACAAAGACTGTCTC 3122  
DB 15718 CAAGATTGTGCACTGCACCTCCAGCTGGGCAACAGACAAAGACTGTCTC 15667  
RESULT 92 AC074136 105956 bp DNA linear PRI 03-APR-2003  
AC074136  
DEFINITION Homo sapiens chromosome 16 clone CTD-231B816, complete sequence.  
ACCESSION AC074136  
VERSION AC074136.4 GI:29501849  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 105956)  
DOE Joint Genome Institute, Stanford Human Genome Center and Los  
Alamos National Laboratory.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 105956)  
DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (15-JUL-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 105956)  
DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (22-AUG-2002) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
4 (bases 1 to 105956)  
DOE Joint Genome Institute, Stanford Human Genome Center and Los  
Alamos National Laboratory.  
TITLE Direct Submission

JOURNAL Submitted (03-APR-2003) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
COMMENT On Apr 3, 2003 this sequence version replaced gi:22417325.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center and Los Alamos  
National Laboratory  
www.sbgc.stanford.edu  
Quality: Phrap Quality >=40 99.9% of Sequence;  
Estimated Total Number of Errors is 0.1.  
FEATURES  
source  
1..105956  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/clone="CTD-231B816"  
ORIGIN  
Query Match 1.7%; Score 52; DB 8; Length 105956;  
Best Local Similarity 100.0%; Pred. No. 5,6e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3071 CAAGATTGTGCACTGCACCTCCAGCTGGGCAACAGACAAAGACTGTCTC 3122  
DB 36426 CAAGATTGTGCACTGCACCTCCAGCTGGGCAACAGACAAAGACTGTCTC 36477  
RESULT 93 AF225899 106539 bp DNA linear PRI 23-FEB-2000  
AF225899/c  
LOCUS Homo sapiens PAC clone 219d7, complete sequence.  
DEFINITION AF225899  
ACCESSION AF225899  
VERSION AF225899.1 GI:7021528  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 106539)  
Fitzpatrick,E.S., Hammond,H.A., DeAngelis,D.M., Soderman,A.R.,  
Wright,J.L., Liu,X., Larson,D., McGowan,J., Ziegler,S.,  
Pritchard,L., Hess,J.F., Todd,J., Caskey,C.T. and Metzger,M.L.  
Direct Submission  
JOURNAL Submitted (18-JAN-2000) Department of Human Genetics, Merck & Co.,  
Inc., Sumneytown Pike, West Point, PA 19486, USA  
TITLE Direct Submission  
JOURNAL Location/Qualifiers  
FEATURES  
source  
1..106539  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="PAC 219d7"  
complement(34..104)  
/rpt\_family="MER5"  
complement(384..662)  
/rpt\_family="Alu"  
1021..1304  
/rpt\_family="Alu"  
1800..2135  
/rpt\_family="Alu"  
2139..2439  
/rpt\_family="Alu"  
3214..3294  
/rpt\_family="MT2B2"  
3588..3657  
/rpt\_family="MT2B2"  
3982..4274  
/rpt\_family="MT2B2"  
5406..5550  
/rpt\_family="MER3"  
5444..5509  
/rpt\_family="MER3"

```
repeat_region complement(5558, .5834)
repeat_region /rpt_family="Alu"
repeat_region /rpt_family="MER3"
repeat_region complement(6534, .6827)
repeat_region /rpt_family="Alu"
repeat_region complement(8239, .8568)
repeat_region /rpt_family="LTR12"
repeat_region 11864, .12159
repeat_region /rpt_family="Alu"
repeat_region complement(12838, .13137)
repeat_region /rpt_family="Alu"
repeat_region complement(14450, .14691)
repeat_region /rpt_family="Alu"
repeat_region complement(15004, .15291)
repeat_region /rpt_family="Alu"
repeat_region complement(16143, .16422)
repeat_region /rpt_family="Alu"
repeat_region complement(16813, .17096)
repeat_region /rpt_family="Alu"
repeat_region complement(17528, .17706)
repeat_region /rpt_family="MIR"
repeat_region 21610, .21890
repeat_region /rpt_family="Alu"
repeat_region 23049, .23329
repeat_region /rpt_family="Alu"
repeat_region 24375, .24585
repeat_region /rpt_family="L1"
repeat_region complement(27993, .28279)
repeat_region /rpt_family="Alu"
repeat_region 28842, .28933
repeat_region /rpt_family="MIR"
STS 30435, .30783
repeat_region /note="LI8270 Human chromosome 2 STS UT426"
repeat_region complement(31317, .31610)
repeat_region /rpt_family="Alu"
repeat_region 32555, .37857
repeat_region /rpt_family="L1"
repeat_region 33985, .34281
repeat_region /rpt_family="Alu"
repeat_region 35580, .35863
repeat_region /rpt_family="Alu"
repeat_region complement(37874, .38708)
repeat_region /rpt_family="L1"
repeat_region 38616, .38712
repeat_region /rpt_family="MIR"
repeat_region complement(41075, .41331)
repeat_region /rpt_family="Alu"
repeat_region complement(42064, .42184)
repeat_region /rpt_family="Alu"
repeat_region 42305, .42413
repeat_region /rpt_family="MIR"
repeat_region complement(42689, .42777)
repeat_region /rpt_family="MIR"
repeat_region 43887, .44156
repeat_region /rpt_family="Alu"
repeat_region 46963, .47240
repeat_region /rpt_family="Alu"
repeat_region complement(47837, .48028)
repeat_region /rpt_family="MIR"
repeat_region complement(51874, .52142)
repeat_region /rpt_family="Alu"
repeat_region 52754, .52801
repeat_region /rpt_family="MER2"
repeat_region 58298, .58529
repeat_region /rpt_family="Alu"
repeat_region 58600, .58721
repeat_region /rpt_family="L1"
repeat_region complement(59586, .59701)
repeat_region /rpt_family="Alu"
repeat_region 60048, .60335
repeat_region /rpt_family="Alu"
repeat_region 60964, .61242

repeat_region /rpt_family="Alu"
repeat_region complement(62514, .63136)
repeat_region /rpt_family="Alu"
repeat_region complement(68228, .68344)
repeat_region /rpt_family="MER20"
STS 69944, .70696
repeat_region /note="G06727 human STS WI-7843"
repeat_region 70776, .71041
repeat_region /rpt_family="Alu"
repeat_region 75420, .75585
repeat_region /rpt_family="Alu"
repeat_region complement(77681, .77951)
repeat_region /rpt_family="Alu"
repeat_region complement(78093, .78393)
repeat_region /rpt_family="Alu"
repeat_region 79977, .80042
repeat_region /rpt_family="THER1"
repeat_region complement(81069, .81363)
repeat_region /rpt_family="Alu"
STS 81561, .81913
repeat_region /note="G09915 human STS CHLC.GGAA19E07.P17317 clone
GGAA19E07"
repeat_region complement(82076, .82380)
repeat_region /rpt_family="Alu"
repeat_region 82595, .82857
repeat_region /rpt_family="Alu"
repeat_region complement(84782, .85059)
repeat_region /rpt_family="Alu"
repeat_region complement(86131, .86655)
repeat_region /rpt_family="Alu"
repeat_region complement(86698, .86772)
repeat_region /rpt_family="MLT1"
repeat_region 87328, .87517
repeat_region /rpt_family="Alu"
repeat_region complement(88025, .88300)
repeat_region /rpt_family="Alu"
repeat_region complement(88382, .88537)
repeat_region /rpt_family="Alu"
repeat_region complement(89103, .92131)
repeat_region /rpt_family="L1"
repeat_region complement(89607, .89870)
repeat_region /rpt_family="Alu"
repeat_region complement(92183, .92468)
repeat_region /rpt_family="Alu"
repeat_region complement(92638, .92723)
repeat_region /rpt_family="L1"
repeat_region complement(93254, .93542)
repeat_region /rpt_family="Alu"
repeat_region 93950, .94195
repeat_region /rpt_family="MER43"
repeat_region 94819, .94882
repeat_region /rpt_family="MER43"
repeat_region 97105, .97472
repeat_region /rpt_family="THER1"
STS 99270, .99669
repeat_region /note="G08815 human STS CHLC.GGAA19E07.P17317 CLONE
GGAA19E07"
repeat_region complement(101199, .101438)
repeat_region /rpt_family="Alu"
repeat_region 101753, .102038
repeat_region /rpt_family="Alu"
repeat_region complement(102216, .102638)
repeat_region /rpt_family="LOR1"
repeat_region complement(103792, .104101)
repeat_region /rpt_family="L1"
repeat_region complement(104170, .104465)
repeat_region /rpt_family="Alu"
repeat_region 104588, .104642
repeat_region /rpt_family="L1"
repeat_region 105260, .105545
repeat_region /rpt_family="Alu"
repeat_region 105891, .106161
repeat_region /rpt_family="Alu"
```

## ORIGIN

Query Match 1.7%; Score 52; DB 8; Length 106539;  
Best Local Similarity 100.0%; Pred. No. 5.6e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122  
DB 12907 CAAGATTGGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 12856

RESULT 94  
BA000025.14/c

## WPCOMMENT

Sequence split into 23 fragments LOCUS BA000025 Accession BA000025

Fragment Name	Begin	End
BA000025_00	1	110000
BA000025_01	100001	210000
BA000025_02	200001	310000
BA000025_03	300001	410000
BA000025_04	400001	510000
BA000025_05	500001	610000
BA000025_06	600001	710000
BA000025_07	700001	810000
BA000025_08	800001	910000
BA000025_09	900001	1010000
BA000025_10	1000001	1110000
BA000025_11	1100001	1210000
BA000025_12	1200001	1310000
BA000025_13	1300001	1410000
BA000025_14	1400001	1510000
BA000025_15	1500001	1610000
BA000025_16	1600001	1710000
BA000025_17	1700001	1810000
BA000025_18	1800001	1910000
BA000025_19	1900001	2010000
BA000025_20	2000001	2110000
BA000025_21	2100001	2210000
BA000025_22	2200001	2229817

Continuation (15 of 23) of BA000025 from base 1400001 (BA000025 Homo sapiens genomic DNA)

Query Match 1.7%; Score 52; DB 8; Length 110000;  
Best Local Similarity 100.0%; Pred. No. 5.6e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122  
DB 29596 CAAGATTGGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 29545

RESULT 95  
BA000041.09/c

## WPCOMMENT

Sequence split into 18 fragments LOCUS BA000041 Accession BA000041

Fragment Name	Begin	End
BA000041_00	1	110000
BA000041_01	100001	210000
BA000041_02	200001	310000
BA000041_03	300001	410000
BA000041_04	400001	510000
BA000041_05	500001	610000
BA000041_06	600001	710000
BA000041_07	700001	810000
BA000041_08	800001	910000
BA000041_09	900001	1010000
BA000041_10	1000001	1110000
BA000041_11	1100001	1210000
BA000041_12	1200001	1310000
BA000041_13	1300001	1410000
BA000041_14	1400001	1510000
BA000041_15	1500001	1610000
BA000041_16	1600001	1710000
BA000041_17	1700001	1750601

Continuation (10 of 18) of BA000041 from base 900001 (BA000041 Pan troglodytes DNA, major contig)

Query Match 1.7%; Score 52; DB 8; Length 110000;  
Best Local Similarity 100.0%; Pred. No. 5.6e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122  
DB 104961 CAAGATTGGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 104910

RESULT 96  
BA000041.10/c

## WPCOMMENT

Sequence split into 18 fragments LOCUS BA000041 Accession BA000041

Fragment Name	Begin	End
BA000041_00	1	110000
BA000041_01	100001	210000
BA000041_02	200001	310000
BA000041_03	300001	410000
BA000041_04	400001	510000
BA000041_05	500001	610000
BA000041_06	600001	710000
BA000041_07	700001	810000
BA000041_08	800001	910000
BA000041_09	900001	1010000
BA000041_10	1000001	1110000
BA000041_11	1100001	1210000
BA000041_12	1200001	1310000
BA000041_13	1300001	1410000
BA000041_14	1400001	1510000
BA000041_15	1500001	1610000
BA000041_16	1600001	1710000
BA000041_17	1700001	1750601

Continuation (11 of 18) of BA000041 from base 1000001 (BA000041 Pan troglodytes DNA, major contig)

Query Match 1.7%; Score 52; DB 8; Length 110000;  
Best Local Similarity 100.0%; Pred. No. 5.6e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122  
DB 4961 CAAGATTGGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 4910

RESULT 97  
HS1164110/c

LOCUS HS1164110 110028 bp DNA linear PRI 18-MAY-2005  
DEFINITION Human DNA sequence from clone RPS-1164110 on chromosome 20q13.13-13.2 Contains part of the BIG2 gene for brefeldin A-inhibited guanine nucleotide-exchange protein 2, ESTs, STSs, GSSs and two CpG islands, complete sequence.

ACCESSION AL049537 GI:9663111

VERSION HTG; BIG2; brefeldin A; CpG island; guanine exchanger.

KEYWORDS Homo sapiens (human)

SOURCE Home sapiens

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 110028)

TITLE Collier R.

JOURNAL Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegase@sanger.ac.uk

COMMENT On Aug 2, 2000 this sequence version replaced gi:9588409. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMBASE; Information on the WORMBASE database can be found at [http://www.sanger.ac.uk/Projects/C\\_elegans/wormpep](http://www.sanger.ac.uk/Projects/C_elegans/wormpep) This sequence was generated from part of bacterial clone contigs of human

chromosome 20, constructed by the Sanger Centre Chromosome 20 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr20>  
RP5-1164110 is from the library RPc1-5 constructed by the group of Plietzer de Jong. For further details see <http://www.choi.org/bacpac/home.htm>  
VECTOR: pcYPAC2  
----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: <http://www.sanger.ac.uk>  
Contact: [vega@sanger.ac.uk](mailto:vega@sanger.ac.uk)  
-----  
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

## FEATURES

source

Location/Qualifiers

misc\_feature

1..110028  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="R2PD:RPc1P7041101164"  
/db\_xref="taxon:9606"  
/chromosome="20"  
/map="q13.13-13.2"  
/clone="RP5-1164110"  
/clone\_1fb="RPc1-5"  
100  
/note="Clone\_right\_end: RP5-906C1"

misc\_feature

1871..1893  
/note="Tandem repeat. Single clone region. Assembly confirmed by restriction enzyme digest data."

misc\_feature

31598..31685  
/note="Single clone region. ALU region. Assembly confirmed by restriction enzyme digest data."

gene

join(32390..32510,51699..51729,52364..52487,61823..61969,63205..63384,64056..64290,74330..74398,76372..76523,79647..79777,81620..81854,82826..82925,83645..83784,85266..85374,86516..86699,95229..95340,95908..96113,98804..98888,98991..99162,99785..99936,100056..100184,101540..101698,104951..105098,106283..106382,108693..108733,108836..109005,AL121903.13:5642..5793,AL121903.13:10804..10976,AL121903.13:12496..12656,AL121903.13:14136..14266,AL121903.13:14403..14532,AL121903.13:16852..16987,AL121903.13:17821..17959,AL121903.13:18068..18122,AL121903.13:19456..19570,AL121903.13:23623..23753,AL121903.13:25885..26053,AL121903.13:29102..29240,AL121903.13:32621..32738,AL121903.13:33595..37265)

/gene="ARFGFP2"

/locus\_tag="RP1-155G6.3-001"

join(32390..32510,51699..51729,52364..52487,61823..61969,63205..63384,64056..64290,74330..74398,76372..76523,79647..79777,81620..81854,82826..82925,83645..83784,85266..85374,86516..86699,95229..95340,95908..96113,98804..98888,98991..99162,99785..99936,100056..100184,101540..101698,104951..105098,106283..106382,108693..108733,108836..109005,AL121903.13:5642..5793,AL121903.13:10804..10976,AL121903.13:12496..12656,AL121903.13:14136..14266,AL121903.13:14403..14532,AL121903.13:16852..16987,AL121903.13:17821..17959,AL121903.13:18068..18122,AL121903.13:19456..19570,AL121903.13:23623..23753,AL121903.13:25885..26053,AL121903.13:29102..29240,AL121903.13:32621..32738,AL121903.13:33595..37265)

/gene="ARFGFP2"

/locus\_tag="RP1-155G6.3-001"

/note="match: ESTs: AA157825.1 AA399434.1 AA628490.1 AA748620.1 AA765985.1 A1425715.1 A1796661.1 A1813340.1 A124958.1 BB864558.1 BF66564.1 BF739950.1 BF909315.1

CDS

B1328134.1 B1467887.1 B1687113.1 B1830616.1 B1847308.1  
BM664817.1 BM826674.1 BM898978.1  
match: CDNA: AF084521.1 AK022623.1 AK074303.1"  
join(32390..32510,51699..51729,52364..52487,61823..61969,63205..63384,64056..64290,74330..74398,76372..76523,79647..79777,81620..81854,82826..82925,83645..83784,85266..85374,86516..86699,95229..95340,95908..96113,98804..98888,98991..99162,99785..99936,100056..100184,101540..101698,104951..105098,106283..106382,108693..108733,108836..109005,AL121903.13:5642..5793,AL121903.13:10804..10976,AL121903.13:12496..12656,AL121903.13:14136..14266,AL121903.13:14403..14532,AL121903.13:16852..16987,AL121903.13:17821..17959,AL121903.13:18068..18122,AL121903.13:19456..19570,AL121903.13:23623..23753,AL121903.13:25885..26053,AL121903.13:29102..29240,AL121903.13:32621..32738,AL121903.13:33595..37271)

/gene="ARFGFP2"

/locus\_tag="RP1-155G6.3-001"

/standard\_name="OTTHMP00000001231"

/note="match: proteins: O46382 O9NKE8 Q9Y6D5 Q9Y6D6"

/codon\_start=1

/protein\_id="CA119320.1"

/db\_xref="GI:56203301"

/translation="MQESQTSMPVSRALBKILADKVRPQHSQRLRACQVLADEIK  
AEIKQRGTAPRANRIBADKYRPERLAQSKSPVSVSTLDCLOKLIYGHITG  
NAPDSGARGKIDRIIVETTSCEFGPPTDESVQOITIKALITANTSPHIEHGTIL  
QVTRCVNIYLAISKNLINQTTAKATLTMTLANTVFRMNQVLAQEARLEKPIQSKEP  
SVTQAAVSPKRVKLSQAOQSKPTPTKEDLTNEHNRDSGKSTENGDAIPRSGS  
LSGTDGDAQEVKDILEDVVTSAIKBAEKHGLTEPRVGLSECOBCAIPGVEN  
SOTNIDANDROSLSDADQAGVAPRSHVLODAFLVFRSLCTLMKPGEG  
PDPKSHIRKRVSLQCLISVQAGVAPRSHVLODAFLVFRSLCTLMKPGEG  
VFRSLATFLTLNFRKMLKQIIVFKKIFANILSTSTSPERHMYIOTLTICA  
DAQCVVDIYVNDYCDLNNANIPELVNDLSKIAQSRSHIELMTPQLDELSKKGIS  
LVSIIKCVNWSKDIYVNHQTSISQGLRITDOEGDGKGLDMARCVTSMSKTVSS  
GTOFTVDDPEQFVYIKQKGLIEHGIILFKPKRGIGLOFLOQMLGTSVVDIQAFL  
HOEBRLDSTONGDFGDSARFKMKVYVYVLODCEKPEFALTEPGLPFGPAQ  
KIDPRMGPARYVEGNGOYLPASADQVLAAGVILIMLTIDHSPOYKMKTKOYI  
KMRGINDSKLPREYLSSTIEBIEGKIMAKETRELITATSTQNVASEKORILY  
NLEMQMKTALMEAVSHAKAPRTSTHLDHVRPMLWTPLLAIVISGLONCD  
TEVSLCEGRCALRIICIFGMDRAVYQALRFLYLAASSGITEKQKIDITIK  
TLIYAHNDGRCALRIICIFGMDRAVYQALRFLYLAASSGITEKQKIDITIK  
LAGEBMKGLNVLVGGVDRKQMSFQSVSTSSVAVVANDRIPTSGTRDGNALV  
DVEFRLCAVSDDELASPHHPRPSLOKIVETSYNNMIRLOMSIMYVIGDHKVG  
CNPNDAIVFVDSLRQKSMKLEKLANRFOQDIPRPHHINKKRSPTIDMAI  
RCIAQMVQSAANIRSGMKNIPAYVHQASDHGNIVELAQVTHIVTTIYQHNHAP  
AIDSFQDVVSCISEPACNAAPDTSMEIRILIRFGKIVSEPRVLQVETSDNNVAP  
GDRVWVGMFPILEFSLCIIRCKLDVTRGILVWFIEIMSKSGHTEPKHMWDLPRIV  
FRIPNMMLPOLSEKSMWTTTCHALYALCDVPTQVYALNEVLISDFVLOQCV  
KODNEQLARSTNCLENIYVINGEKSEGVWDETCMLDIKFTTIPHYLTWRVGM  
EDDSERKLVDLDKQSLSSIDKPSERKQSLSPNTDSKRGKRYANOKFLASLLIK  
CVVDLELQTDIDNIVFPATSKEDAEHMAVAAQDTLADHIFREDQMYVYMSQH  
LFKLDLCQESHSFSAKNSNYEQRTVAVRQKESKPNLKKQTSLSACLRILKDE  
MYVDENRDSWBERIOQRLLTVCEBALAFITVNSSHREAMTSLLLTLTKLKNDE  
KPKHANSYVYLCIMQFDLIPBLRAVLRKFLRIGVVYKIMWIEPESQVPAALSPV  
W"

49711

/gene="ARFGFP2"

/locus\_tag="RP1-155G6.3-001"

/note="Clone\_left\_end: RP1-155G6"  
69981  
/gene="ARFGFP2"  
/locus\_tag="RP1-155G6.3-001"  
/note="Tandem repeat. Forced join. Gap size estimated to be approximately 360bp by BamHI digest data and pUC bridge sizing."  
71108..71162  
/gene="ARFGFP2"  
/locus\_tag="RP1-155G6.3-001"  
/note="Single clone region. ALU region. pUC subcloned to give averaged 20x coverage. Assembly confirmed by restriction enzyme digest data."

misc\_feature  
misc\_feature  
misc\_feature  
misc\_feature  
misc\_feature

misc\_feature

```
/gene="ARFGSP2"
/locus_tag="RP1-155G6.3-001"
/note="Single clone region. ALU region. pUC subcloned to
give average 4x coverage. Assembly confirmed by
restriction enzyme digest data."
join(108600..108733,108836..108946)
/gene="ARFGSP2"
/locus_tag="RP1-155G6.3-002"
join(108600..108733,108836..108946)
/gene="ARFGSP2"
/locus_tag="RP1-155G6.3-002"
join(108600..108733,108836..108946)
/misc_feature
/note="match: ESTs: BF83823.1"
110028
/note="Clone_right_end: RP5-1164110"

ORIGIN
Query Match      1.7%; Score 52; DB 8; Length 110028;
Best Local Similarity 100.0%; Pred. No. 5.6e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3071 CAAGATTGTCACCTGCTCCAGCTGCGCAACAGAGCAAGACTCTGTCTC 3122
        |||
Db      6165 CAAGATTGTCACCTGCTCCAGCTGCGCAACAGAGCAAGACTCTGTCTC 6114

RESULT 98
AL590788      110042 bp      DNA      linear      PRI 18-MAY-2005
LOCUS      Human DNA sequence from clone RP1-210P2 on chromosome 6, complete
ACCESSION      AL590788
VERSION      AL590788
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE      1 (bases 1 to 110042)
AUTHORS      Clark,S.
TITLE      Direct Submision
JOURNAL      Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
COMMENTS      Clone requests: clonerequests@sanger.ac.uk
On Jun 14, 2001 this sequence version replaced gi:1418913.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormep
This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Ch6
RP1-210P2 is from the library RPCI-11.1 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBACE3.6
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
Location/Qualifiers
1..110042

/gene="ARFGSP2"
/locus_tag="RP1-155G6.3-001"
/note="Single clone region. ALU region. pUC subcloned to
give average 4x coverage. Assembly confirmed by
restriction enzyme digest data."
join(108600..108733,108836..108946)
/gene="ARFGSP2"
/locus_tag="RP1-155G6.3-002"
join(108600..108733,108836..108946)
/gene="ARFGSP2"
/locus_tag="RP1-155G6.3-002"
join(108600..108733,108836..108946)
/misc_feature
/note="match: ESTs: BF83823.1"
110028
/note="Clone_right_end: RP5-1164110"

ORIGIN
Query Match      1.7%; Score 52; DB 8; Length 110042;
Best Local Similarity 100.0%; Pred. No. 5.6e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2889 GAGCAGGTGATGATCACTGAGCCAGAGTTCGAGACCGCTGCGCAACAT 2940
        |||
Db      63836 GAGCAGGTGATGATCACTGAGCCAGAGTTCGAGACCGCTGCGCAACAT 63887

RESULT 99
AF322449/c
LOCUS      Homo sapiens chromosome 17 clone BAC347P24 map 17p13.3, ***
DEFINITION      SEQUENCING IN PROGRESS ***, 26 unordered pieces.
ACCESSION      AF322449
VERSION      AF322449.1 GI:11559854
KEYWORDS      HTG, HTGS, PHASE1
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE      1 (bases 1 to 11051)
AUTHORS      Zhao,X.T., He,M., Wan,D.F., Ye,Y., Qin,W.X., Huang,Y., Zuo,L. and
Gu,J.R.
TITLE      Gene clone on human chromosome 17p13.3
JOURNAL      Unpublished
AUTHORS      2 (bases 1 to 11051)
Zhao,X.T., He,M., Wan,D.F., Ye,Y., Qin,W.X., Huang,Y., Zuo,L. and
Gu,J.R.
TITLE      Direct Submision
JOURNAL      Submitted (21-NOV-2000) National Laboratory For Oncogenes & Related
Genes, Shanghai Cancer Institute, 25/ln 2200 Xie-Tu Road, Shanghai
200032, P. R. China
COMMENTS      * NOTE: This is a 'working draft' sequence. It currently
* consists of 26 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1      484: contig of 484 bp in length
*      485      gap of unknown length
*      485      1176: contig of 692 bp in length
*      1177      gap of unknown length
*      1177      1450: contig of 274 bp in length
*      1451      gap of unknown length
*      1451      2440: contig of 990 bp in length
*      2441      gap of unknown length
*      2441      3647: contig of 1207 bp in length
*      3648      gap of unknown length
*      4231      5202: contig of 583 bp in length
*      5202      gap of unknown length
*      5202      6225: contig of 972 bp in length
*      6225      gap of unknown length
*      6225      1023: contig of 1023 bp in length
*      1023      gap of unknown length
```





566	<800	5842	6097	510	<800
10138	10241	726	722	6511	6403
5932	6176	8963	9441	1997	1982
3547	3470	22693	22397	22133	22287
5667	5909	9245	9441	5286	5248
3422	3470	6720	7134	5832	5854
12502	12708	3489	3562	1444	1430
9097	9443	15829	16232	7147	7100
3849	3878	3900	4061	6905	7100
9736	10241	81	<800	2371	2343
3368	3470	12595	12874	4605	4618
128	<800	10029	10085	12762	13005
6610	6926	2887	2925	3492	3534
510	<800	334	<800	5177	5248
15047	16041	2471	2532	3147	3118
18116	18484	9535	10085	3049	3118
				4670	4618
				13645	13005
				7458	7605

## FEATURES

## source

Location/Qualifiers

1.11688  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clone\_id="RP13-487A18"  
/clone\_11b="null"  
1.5  
/note="Single subclone region"  
116885.116888  
/note="Single subclone region"

## ORIGIN

Query Match 1.7%; Score 52; DB 8; Length 116888;  
Best Local Similarity 100.0%; Pred. No. 5.6e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCGAGTGATCACTGAGGCCAGAGTTGAGACCGAGCTGGCCAACT 2940  
|||||  
DB 26574 GAGGCGAGTGATCACTGAGGCCAGAGTTGAGACCGAGCTGGCCAACT 26523

## RESULT 101

AC016590 117899 bp DNA linear PRI 22-MAR-2003  
LOCUS Homo sapiens chromosome 19 clone CTD-3220F14, complete sequence.  
AC016590  
VERSION AC016590.7 GI:28191395  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE 1 (bases 1 to 117899)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 117899)  
AUTHORS DOE Joint Genome Institute.  
TITLE Direct Submission  
JOURNAL Submitted (04-DEC-1999) Production Sequencing Facility, DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE 3 (bases 1 to 117899)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (01-FEB-2003) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
REFERENCE 4 (bases 1 to 117899)  
AUTHORS DOE Joint Genome Institute and Stanford Human Genome Center.  
TITLE Direct Submission  
JOURNAL Submitted (22-MAR-2003) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
COMMENT On Feb 1, 2003 this sequence version replaced gi:13695950.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 99.5% of Sequence;  
Estimated Total Number of Errors is 0.6  
NOTE: Large tandem repeat 57800-92700. There are possibly more repeat copies than are represented in this assembly.  
NOTE: This insert is not the entire sequence of the clone (entire sequence is 200kb). It is clipped at the overlaps with AC008806 and AC012309. The number of bases overlapped with AC008806 is 7498 bps and with AC012309 is 13284 bps.

## FEATURES

## source

Location/Qualifiers

1.117899  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="19"  
/clone="CTD-3220F14"  
57800..92700  
/note="NOTE: Large tandem repeat 57800-92700. There are possibly more repeat copies than are represented in this assembly."

## ORIGIN

Query Match 1.7%; Score 52; DB 8; Length 117899;  
Best Local Similarity 100.0%; Pred. No. 5.6e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCGAGTGATCACTGAGGCCAGAGTTGAGACCGAGCTGGCCAACT 2940  
|||||  
DB 8892 GAGGCGAGTGATCACTGAGGCCAGAGTTGAGACCGAGCTGGCCAACT 8943

## RESULT 102

AC104051/c 120723 bp DNA linear PRI 02-MAR-2002  
LOCUS Homo sapiens chromosome 8, clone CTD-2339F6, complete sequence.  
AC104051  
VERSION AC104051.3 GI:19071664  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homniidae; Homo.  
REFERENCE 1 (bases 1 to 120723)  
AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
TITLE Homo sapiens chromosome 8, clone CTD-2339F6  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 120723)  
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N.,

Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazarro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Labèque, K., Lamazares, R., Landers, T., Lehoczeky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheters, R., Meldrim, J., Menues, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhng, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severly, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggilo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

**TITLE**  
Direct Submission  
**JOURNAL**  
Submitted (01-DEC-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
**REFERENCE**  
3 (bases 1 to 120723)  
**AUTHORS**  
Birren, B., Linton, L., Nusbaum, C., Lander, B., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouhgalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazarro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Labèque, K., Lamazares, R., Landers, T., Lehoczeky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margulis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., Meldrim, J., Menues, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhng, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schnupack, R., Seaman, S., Severly, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Triggilo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

**TITLE**  
Direct Submission  
**JOURNAL**  
Submitted (02-MAR-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
**COMMENT**  
On Mar 2, 2002 this sequence version replaced g1:18450132.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

**FEATURES**  
**source**  
-----  
Location/Qualifiers  
1. 120723  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"  
/clone="CTD-2339F6"  
/clone\_lib="CITD1 Human BAC"  
repeat\_region  
35. 813

/rpt\_family="LIM4C"  
857. 1128  
/rpt\_family="LIM4C"  
1183. 1373  
/rpt\_family="LIM2"  
1767. 1925  
/rpt\_family="L2"  
2029. 2076  
/rpt\_family="AT\_rich"  
2097. 2120  
/rpt\_family="AT\_rich"  
complement(3952. 4083)  
/rpt\_family="MERSA"  
complement(4837. 4932)  
/rpt\_family="MIR"  
complement(6143. 6287)  
/rpt\_family="MIR"  
6340. 6709  
/rpt\_family="MSTB"  
complement(6710. 7011)  
/rpt\_family="ALUSX"  
7012. 7052  
/rpt\_family="MSTB"  
complement(8754. 8908)  
/rpt\_family="MLT1L"  
complement(9576. 9818)  
/rpt\_family="LIPAB"  
complement(9819. 11131)  
/rpt\_family="LIPAB"  
complement(11132. 11413)  
/rpt\_family="ALUSP"  
complement(11414. 15347)  
/rpt\_family="LIPAB"  
15504. 15552  
/rpt\_family="L2"  
complement(16682. 17158)  
/rpt\_family="LRR33"  
17190. 17237  
/rpt\_family="TG)n"  
17266. 17289  
/rpt\_family="AT\_rich"  
complement(19324. 19610)  
/rpt\_family="ALUSX"  
20277. 20311  
/rpt\_family="AT\_rich"  
21092. 21151  
/rpt\_family="CA)n"  
22789. 23255  
/rpt\_family="L2"  
complement(23674. 25756)  
/rpt\_family="LIMB7"  
complement(25757. 25934)  
/rpt\_family="LIM4C"  
complement(25935. 26325)  
/rpt\_family="MLT1B"  
complement(26326. 27976)  
/rpt\_family="LIM4C"  
complement(27977. 28340)  
/rpt\_family="MLT1A1"  
complement(28341. 28642)  
/rpt\_family="LIM4C"  
28647. 28700  
/rpt\_family="TG)n"  
complement(30192. 30366)  
/rpt\_family="ALUY"  
complement(30367. 33534)  
/rpt\_family="LIPAB3"  
complement(33535. 33656)  
/rpt\_family="ALUY"  
33967. 34276  
/rpt\_family="ALUSX"  
37011. 37040  
/rpt\_family="AT\_rich"



AUTHORS  
TITLE  
JOURNAL

## COMMENT

Tracey, A.  
Direct Submission  
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,  
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk  
Clone requests: clonerequest@sanger.ac.uk  
On Jul 6, 2001 this sequence version replaced gi:14575385.  
The following abbreviations are used to associate primary accession  
numbers given in the feature table with their source databases:  
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WORMPEP; Information  
on the WORMPEP database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormep This sequence  
was generated from part of bacterial clone contigs of human  
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping  
Group. Further information can be found at  
http://www.sanger.ac.uk/HGP/Chr6  
RP11-41204 is from the library RPCT-11.2 constructed by the group  
of Pieter de Jong. For further details see  
http://www.chori.org/bacpac/home.htm  
VECTOR: pBACe3.6

----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: vegas@sanger.ac.uk  
-----

This sequence was finished as follows unless otherwise noted: all  
regions were either double-stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one subclone; and the assembly was confirmed by restriction digest,  
except on the rare occasion of the clone being a YAC.

## FEATURES

## source

Location/Qualifiers

1. 130146

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="R2PD:RPCIB753004412"

/db\_xref="taxon:9606"

/chromosome="6"

/map="q16.2-21"

/clone="RP11-41204"

/clone\_1fb="RPCT-11.2"

## misc\_feature

1

/note="Clone left end: RP11-41204"

join(AL121957.15:80152..80726,complement(77547..77681),

complement(17378..17576),complement(7535..7861))

/gene="C6orf210"

/locus\_tag="RP11-5919.3-002"

join(AL121957.15:80152..80726,complement(77547..77681),

complement(17378..17576),complement(7535..7861))

/gene="C6orf210"

/locus\_tag="RP11-5919.3-002"

complement(7537)

/gene="C6orf210"

/locus\_tag="RP11-5919.3-002"

complement(7549..7554)

/gene="C6orf210"

/locus\_tag="RP11-5919.3-002"

join(AL121957.15:80431..80726,complement(77547..77681),

complement(17378..17576),complement(7769..7861))

/gene="C6orf210"

/locus\_tag="RP11-5919.3-002"

/standard\_name="OTTHUMP000006565"

/codon\_start=1

/product="chromosome 6 open reading frame 210"

/protein\_id="CAH73572.1"

/db\_xref="GI:5565806"

/db\_xref="GOA:Q5VU39"

## gene

/db\_xref="UniProt/TrEMBL:Q5VU39"  
/translation="MNFRLQLHLPRYVIGAGSGSPRLWMSPLDITISSVSGSRGSSK  
SPAHNVQVSAEKIVGPTFSFMSRLCLISBELSIAQVRLVGTQHPLLTAGGLV  
HDSNMSLDRGLVLLISKAAGPSVNSCONVDSIYSGQRLAETTELTHALL  
VHRGVINELQSSDPKDMQFGNKIILISGDFLAAACGALLQNTKSFENGPI  
AIYQGDSESMWILSKHPRLS"  
join(AL121957.15:80153..80726,complement(77547..77681),  
complement(17378..17576),complement(7549..7554),  
complement(17378..17576),complement(7537..7549),  
complement(17378..17576),complement(7535..7861))

## mRNA

join(AL121957.15:80153..80726,complement(77547..77681),  
complement(17378..17576),complement(7549..7554),  
complement(17378..17576),complement(7537..7549),  
complement(17378..17576),complement(7535..7861))

## CDS

/product="chromosome 6 open reading frame 210"  
/note="match: BSTR: AA007421 AA007488 AA633846 AI033207  
AI097731 A1492248 A1521803 A1678112 A1765825 A1796241  
A1827607 AM243064 AM276704 AM408635 AM575740 AM619911  
BF099861 BB101041 BE221188 BE283879 BE617435 BE818989  
BF509925 BG292722 BG401578 B1180782 B1465307 B1551760  
BJ002337 BJ006031 BJ064031 BM017975 BA488969 BM796336  
BM82264 BM836490 BM974205 H49955 H49952  
match: CDNA: AF254956 AF254956.1 AJ420575.1 AK000293  
AK017329"  
join(AL121957.15:80431..80726,complement(77547..77681),  
complement(17378..17576),complement(7549..7554),  
complement(17378..17576),complement(7537..7549),  
complement(17378..17576),complement(7535..7861))

/product="chromosome 6 open reading frame 210"  
/protein\_id="CAH73573.1"  
/db\_xref="GI:5565807"  
/db\_xref="GOA:Q5VU38"  
/db\_xref="UniProt/TrEMBL:Q5VU38"  
/translation="MNFRLQLHLPRYVIGAGSGSPRLWMSPLDITISSVSGSRGSSK  
SPAHNVQVSAEKIVGPTFSFMSRLCLISBELSIAQVRLVGTQHPLLTAGGLV  
HDSNMSLDRGLVLLISKAAGPSVNSCONVDSIYSGQRLAETTELTHALL  
VHRGVINELQSSDPKDMQFGNKIILISGDFLAAACGALLQNTKSFENGPI  
AIYQGDSESMWILSKHPRLS"  
join(AL121957.15:80431..80726,complement(77547..77681),  
complement(17378..17576),complement(7549..7554),  
complement(17378..17576),complement(7537..7549),  
complement(17378..17576),complement(7535..7861))

## misc\_feature

72128

/note="Clone right end: RP11-300D11"

## misc\_feature

128147

/note="Clone left end: RP1-67A8"

## ORIGIN

Query Match 1.7% Score 52; DB 8; Length 130146;  
Best Local Similarity 100.0%; Pred. No. 5.6e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 3071 CAAGTTGTGCGACGCTCTGAGCGGCAACGACCAAGACTCTGCTC 3122  
|||||  
DB 69521 CAAGTTGTGCGACGCTCTGAGCGGCAACGACCAAGACTCTGCTC 69470  
|||||

RESULT 105  
AC024096 130572 bp DNA 1linear HTG 08-JAN-2003  
LOCUS Homo sapiens chromosome 3 clone RP11-214A22, WORKING DRAFT  
DEFINITION AC024096  
SEQUENCE AC024096  
AC024096.16 GI:20335482  
VERSION HTG, HTGS\_PHASE1, HTGS\_DRAFT.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 130572)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Alsbrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T., Barbiste,J., Benton,J., Benton,J., Bimaga,K., Blankenburg,K., Bonnin,D., Bouck,J., Bowls,S., Brieva,M., Brown,E., Brown,N., Bryant,N.P., Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabriel,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,M., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogues,M., Holloway,C., Hollins,B., Homes,F., Howard,S., Huber,J., Hulyk,S., Hume,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J., Kovar,C., Kratovic,J., Kureishi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Lien,C., Liu,J., Liu,M., Louleghed,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mawhinney,B., McLeod,M.P., Meador,M., Mei,G., Metzger,M., Miner,G., Miner,Z., Mitchell,T., Mohabat,K., Morgan,M., Morris,S., Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N., Nickerson,E., Nwankwo,S., Oguh,M., Okunuga,G., Oragunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Pickens,R., Prims,E., Pu,L.L., Quiles,M., Ren,Y., Rivers,M., Rojas,A., Rojudoan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shooshari,N., Sisson,I., Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A., Tabor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Umanai,K., Vazquez,L., Vera,V., Villalón,D., Vinson,R., Wang,O., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Wallington,S., Williams,G., Williamson,A., Wleczkyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D., Weinstein,G., and Gibbs,R.  
Direct Submission  
2 (bases 1 to 130572)  
Worley,K.C.  
Direct Submission  
Submitted (24-FEB-2000) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 130572)  
Worley,K.C.  
Direct Submission  
Submitted (08-JAN-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
COMMENT  
----- Genome Center  
On Apr 28, 2002 this sequence version replaced gi:16117892.  
Center: Baylor College of Medicine  
Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: HXG  
Center clone name: RP11-214A22  
----- Summary Statistics  
Sequencing vector: M13;  
Chemistry: Dye-terminator Big Dye; 95% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 106061 bases at least Q40  
Consensus quality: 115099 bases at least Q30  
Consensus quality: 121518 bases at least Q20  
Estimated insert size: 143263; sum-of-contigs estimation  
Quality coverage: 4x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 20 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1  
2637 2636: contig of 2636 bp in length  
2737 2736: gap of unknown length  
5887 5887: contig of 3151 bp in length  
5988 5987: gap of unknown length  
8254 8254: contig of 2267 bp in length  
8255 8254: gap of unknown length  
8355 10964: contig of 2610 bp in length  
10965 11064: gap of unknown length  
11065 13703: contig of 2639 bp in length  
13704 13803: gap of unknown length  
16745 16745: contig of 2942 bp in length  
16746 16845: gap of unknown length  
16846 20065: contig of 3221 bp in length  
20067 20166: gap of unknown length  
25167 25320: contig of 5054 bp in length  
25321 25320: gap of unknown length  
29845 29845: contig of 4525 bp in length  
29846 29945: gap of unknown length  
35872 35872: contig of 5927 bp in length  
35873 35973: gap of unknown length  
35973 35973: contig of 3710 bp in length  
39783 39782: gap of unknown length  
39783 39782: gap of unknown length  
44248 44248: contig of 4466 bp in length  
44249 44348: gap of unknown length  
50564 50564: contig of 6216 bp in length  
50565 50664: gap of unknown length  
59327 59327: contig of 8663 bp in length  
59328 59427: gap of unknown length  
59428 59427: gap of unknown length  
66662 66662: contig of 7235 bp in length  
66663 66762: gap of unknown length  
74213 74213: contig of 7451 bp in length  
74214 74313: gap of unknown length  
74314 86510: contig of 12197 bp in length  
86511 86610: gap of unknown length  
86611 102506: contig of 15956 bp in length  
102507 102606: gap of unknown length  
102607 115136: contig of 12530 bp in length  
115137 115236: gap of unknown length  
115237 130572: contig of 13536 bp in length.  
Location/Qualifiers  
1..130572  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/clone="RP11-214A22"  
2637..2736  
gap

```

gap      /estimated_length=unknown
          5888..5987
          /estimated_length=unknown
gap      9255..8354
          /estimated_length=unknown
gap      10965..11064
          /estimated_length=unknown
gap      13704..13803
          /estimated_length=unknown
gap      16746..16845
          /estimated_length=unknown
gap      20067..20166
          /estimated_length=unknown
gap      25221..25320
          /estimated_length=unknown
gap      29846..29945
          /estimated_length=unknown
gap      35873..35972
          /estimated_length=unknown
gap      39683..39782
          /estimated_length=unknown
gap      44249..44348
          /estimated_length=unknown
gap      50565..50664
          /estimated_length=unknown
gap      59328..59427
          /estimated_length=unknown
gap      66663..66762
          /estimated_length=unknown
gap      74214..74313
          /estimated_length=unknown
gap      86511..86610
          /estimated_length=unknown
gap      102507..102606
          /estimated_length=unknown
gap      115137..115236
          /estimated_length=unknown

```

## ORIGIN

```

Query Match      1.7%; Score 52; DB 14; Length 130572;
Best Local Similarity 100.0%; Pred.No. 5.6e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy      3071 CAAGATTGCGACATGACACTCCAGCGCTGGGCAAGAGACAGACTCTGTCTC 3122
Db      113714 CAAGATTGCGACATGACACTCCAGCGCTGGGCAAGAGACAGACTCTGTCTC 113765

```

```

RESULT 106
AC022766      130984 bp      DNA      linear      HTG 01-JUN-2000
LOCUS      Homo sapiens clone RP11-318E23, WORKING DRAFT SEQUENCE, 31
DEFINITION      Unpublished.
ACCESSION      AC022766
VERSION      AC022766.3 GI:8138013
KEYWORDS      HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 130984)
Birtren,B., Linton,L., Nusbaum,C. and Lander,B.
Homo sapiens, clone RP11-318E23
Unpublished
2 (bases 1 to 130984)
Birtren,B., Linton,L., Nusbaum,C., Lander,B., Adnan,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Beda,F.,
Boguslavsky,L., Bouhgalter,B., Brown,A., Burkett,G., Caele,A.,
Chapel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
DeRellano,K., Dewar,K., Domino,M., Doyle,M., Feneator,J.,
Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagoe,B., Heaford,A., Horton,L.,

```

TITLE  
JOURNAL  
COMMENT

```

Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,U.,
Landers,T., Lehoczy,J., Levine,R., Lien,C., Liu,G., Locke,K.,
Macdonald,P., Marcus,N., McEwan,P., McNair,A., McKernan,K.,
McPherson,R., Meldrum,J., Meneses,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Leary,T.M., Peterson,K.,
Pierre,N., Pisan,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tefaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jun 1, 2000 this sequence version replaced gi:7712161.
All repeats were identified using RepeatMasker:
Smit, A.P. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RW/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: 15852
Center clone name: 318_E_23
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 118950 bases at least Q40
Consensus quality: 124428 bases at least Q30
Consensus quality: 126573 bases at least Q20
Insert size: 152000; agarose-gel
Insert size: 127984; sum-of-contigs
Quality coverage: 4.6 in Q20 bases; agarose-gel
Quality coverage: 5.5 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 31 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1      1066: contig of 1066 bp in length
*      1067      1166: gap of 100 bp
*      1167      2747: contig of 1581 bp in length
*      2748      2847: gap of 100 bp
*      2848      4584: contig of 1737 bp in length
*      4585      4685: gap of 100 bp
*      4686      6489: contig of 1805 bp in length
*      6490      6590: gap of 100 bp
*      6591      8380: contig of 1791 bp in length
*      8381      8480: gap of 100 bp
*      8481      10218: contig of 1738 bp in length
*      10219      10318: gap of 100 bp
*      10319      12338: contig of 2020 bp in length
*      12339      12438: gap of 100 bp
*      12439      14278: contig of 1840 bp in length
*      14279      14378: gap of 100 bp
*      14379      16225: contig of 1847 bp in length
*      16226      16325: gap of 100 bp
*      16326      18627: contig of 2302 bp in length
*      18628      18727: gap of 100 bp
*      18728      21450: contig of 2723 bp in length
*      21451      21550: gap of 100 bp
*      21551      25430: contig of 3880 bp in length
*      25431      25530: gap of 100 bp
*      25531      28870: contig of 3340 bp in length
*      28871      28971: gap of 100 bp
*      28972      32940: contig of 3370 bp in length
*      32941      33040: gap of 100 bp

```



```

* 33041 35975: contig of 2935 bp in length
* 35976 36075: gap of 100 bp
* 36076 39460: contig of 3385 bp in length
* 39461 39560: gap of 100 bp
* 39561 43197: contig of 3637 bp in length
* 43198 43297: gap of 100 bp
* 43298 46333: contig of 3036 bp in length
* 46334 46433: gap of 100 bp
* 46434 51625: contig of 5182 bp in length
* 51626 51725: gap of 100 bp
* 51726 55173: contig of 3448 bp in length
* 55174 55273: gap of 100 bp
* 55274 59400: contig of 4127 bp in length
* 59401 59500: gap of 100 bp
* 59501 64607: contig of 5307 bp in length
* 64608 64907: gap of 100 bp
* 64908 68668: contig of 3761 bp in length
* 68669 68768: gap of 100 bp
* 68769 72524: contig of 3756 bp in length
* 72525 77709: contig of 5085 bp in length
* 77710 77809: gap of 100 bp
* 77810 83970: contig of 6161 bp in length
* 83971 84070: gap of 100 bp
* 84071 89106: contig of 5036 bp in length
* 89107 89206: gap of 100 bp
* 89207 96647: contig of 7411 bp in length
* 96648 96747: gap of 100 bp
* 96748 105147: contig of 8400 bp in length
* 105148 105247: gap of 100 bp
* 105248 118840: contig of 13593 bp in length
* 118841 118941: gap of 100 bp
* 118941 130984: contig of 12044 bp in length.
Location/Qualifiers
1..130984

```

```

FEATURES
source
1..130984
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-318523"
/clone_1fb="RP11 Human Male BAC"
misc_feature
1..1066
/note="assembly_fragment"
gap
1067..1166
/estimated_length=100
misc_feature
1167..2747
/note="assembly_fragment"
gap
2748..2847
/estimated_length=100
misc_feature
2848..4584
/note="assembly_fragment"
gap
4585..4684
/estimated_length=100
misc_feature
4685..6489
/note="assembly_fragment"
gap
6490..6589
/estimated_length=100
misc_feature
6590..8380
/note="assembly_fragment"
gap
8381..8480
/estimated_length=100
misc_feature
8481..10218
/note="assembly_fragment"
gap
10219..10318
/estimated_length=100
misc_feature
10319..12338
/note="assembly_fragment"
gap
12339..12438
/estimated_length=100
misc_feature
12439..14278
/note="assembly_fragment"
gap
14279..14378
/estimated_length=100
misc_feature
14379..16225

```

```

/note="assembly_fragment"
16226..16325
/estimated_length=100
misc_feature
16326..18627
/note="assembly_fragment"
18628..18727
/estimated_length=100
misc_feature
18728..21450
/note="assembly_fragment"
21451..21550
/estimated_length=100
misc_feature
21551..25430
/note="assembly_fragment"
25431..25530
/estimated_length=100
misc_feature
25531..28870
/note="assembly_fragment"
28871..28970
/estimated_length=100
misc_feature
28971..32540
/note="assembly_fragment"
32541..33040
/estimated_length=100
misc_feature
33041..35975
/note="assembly_fragment"
35976..36075
/estimated_length=100
misc_feature
36076..39460
/note="assembly_fragment"
39461..39560
/estimated_length=100

```

Query Match 1.74; Score 52; DB 14; Length 130984;  
 Best Local Similarity 100.0%; Pred. No. 5; Ce-16;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCACCTGCACCTGCAGCTGGGCAAGCAAGACTGTCTC 3122  
 Db 18564 CAAATTGTGCACCTGCACCTGCAGCTGGGCAAGCAAGACTGTCTC 18615

RESULT 107  
 AC100793/c 134465 bp DNA linear PRI 15-JAN-2003  
 LOCUS  
 DEFINITION Homo sapiens chromosome 17, clone CTD-3193K9, complete sequence.  
 ACCESSION AC100793  
 VERSION AC100793.8 GI:2753789  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 1 (bases 1 to 134465)  
 Birren,B., Nusbaum,C. and Lander,E.  
 Homo sapiens chromosome 17, clone CTD-3193K9  
 Unpublished  
 2 (bases 1 to 134465)  
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
 Anderson,S., Barris,N., Baethien,V., Bogunlavsky,L., Boukhalter,B.,  
 Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,  
 Choepe,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A.,  
 Cooke,P., Dearrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,  
 Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,  
 Ginde,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
 Hagos,B., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R.,  
 Jones,C., Kamat,A., Karatas,A., Kells,C., Lacroque,K.,  
 Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G.,  
 Maclean,C., Macdonald,P., Major,J., Marguis,N., Matthews,C.,  
 McCarthy,M., McEwan,P., McKernan,K., McNetters,R., Meldrum,J.,  
 Menue,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,  
 Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D.,  
 Oliver,J., Peterson,K., Phunhkhang,P., Pierre,N., Pollara,V.,

Raymond,C., Retta,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Straus,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE  
JOURNAL

Submitted (22-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE  
AUTHORS

3 (bases 1 to 134465)  
Birken,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barra,N., Baslein,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., MacDonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Menue,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE  
JOURNAL

Submitted (27-OCT-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

REFERENCE  
AUTHORS

4 (bases 1 to 134465)  
Birken,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barra,N., Baslein,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeArrellano,K., Dewar,K., Diaz,J.S., Dodge,S., Fato,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N., Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., Maclean,C., MacDonald,P., Major,J., Matthews,C., McCarthy,M., Meldrim,J., Menue,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE  
JOURNAL

Submitted (15-JAN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

COMMENT

On Jan 15, 2003 this sequence version replaced g124415443.

All repeats were identified using RepeatMasker:

ftp://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIMR

Web site: <http://www-seg.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

Project Information

Center project name: L20710

Center clone name: 3193\_X\_9

FEATURES

source

Location/Qualifiers  
1..134465  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="17"

/map="17"  
/clone\_lib="CTD-3193X9"  
/clone\_lib="CTD2 Human BAC"  
1..88  
/rpt\_family="AluUo/FRAM"  
94..389  
/rpt\_family="AluSx"  
complement(479..551)  
/rpt\_family="Tigerg5"  
575..694  
/rpt\_family="MER21B"  
1296..1431  
/rpt\_family="AluUb"  
1432..1728  
/rpt\_family="AluSx"  
1731..2062  
/rpt\_family="AluUb"  
2400..2547  
/rpt\_family="AluSx"  
2613..2661  
/rpt\_family="AluSx"  
complement(2670..2789)  
/rpt\_family="FLAM\_C"  
complement(2795..3122)  
/rpt\_family="AluY"  
6180..6232  
/rpt\_family="AluSx"  
complement(6233..6530)  
/rpt\_family="AluY"  
6531..6585  
/rpt\_family="AluSx"  
complement(6586..6887)  
/rpt\_family="AluSx"  
6888..7159  
/rpt\_family="AluSx"  
7875..7912  
/rpt\_family="AluSx"  
7957..8044  
/rpt\_family="AluSx"  
9590..9691  
/rpt\_family="MER91A"  
10084..10105  
/rpt\_family="AluSx"  
10488..10589  
/rpt\_family="AluSx"  
12437..12588  
/rpt\_family="MIR3"  
12868..12900  
/rpt\_family="AluSx"  
13333..13370  
/rpt\_family="AluSx"  
14420..14475  
/rpt\_family="AluSx"  
14721..14744  
/rpt\_family="AluSx"  
14801..14840  
/rpt\_family="AluSx"  
15881..15924  
/rpt\_family="AluSx"  
16110..16162  
/rpt\_family="AluSx"  
complement(16380..16658)  
/rpt\_family="AluSx"  
complement(16694..16909)  
/rpt\_family="AluSx"  
17073..17108  
/rpt\_family="AluSx"  
17378..17409  
/rpt\_family="AluSx"  
18578..18710  
/rpt\_family="MIR"

```
repeat_region complement(18775..19088)
/rpt_family="AluSc"
repeat_region 19780..19886
/rpt_family="GA-rich"
repeat_region 19935..20031
/rpt_family="MIR"
repeat_region 20337..20445
/rpt_family="GA-rich"
repeat_region 20855..20960
/rpt_family="(CTGGGG)n"
repeat_region 20906..21085
/rpt_family="(TTGGGG)n"
repeat_region 21037..21216
/rpt_family="(CTGGGG)n"
repeat_region 21166..21309
/rpt_family="(TTGGGG)n"
repeat_region complement(22213..22416)
/rpt_family="L2"
repeat_region 22417..22720
/rpt_family="AluSc"
repeat_region 23332..23457
/rpt_family="MIR"
repeat_region complement(23477..23601)
/rpt_family="MIR"
```

Query Match 1.7%; Score 52; DB 8; Length 134465;  
Best Local Similarity 100.0%; Pred. No. 5.6e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3071 CAAGATTGTCACGTCGACCTCGGCGCAAGACAGACAGACTGTCTC 3122

45103 CAAAGTTGTCACGTCGACCTCGGCGCAAGACAGACAGACTGTCTC 45052

RESULT 108  
AC069511  
LOCUS AC069511 135805 bp DNA linear HTG 25-JUL-2000  
DEFINITION Homo sapiens chromosome 3 clone RP11-224122, WORKING DRAFT  
ACCESSION AC069511  
VERSION AC069511.4 GI:9438659  
KEYWORDS HTG; HTGS PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE  
AUTHORS 1 (bases 1 to 135805)  
Muzny,D.M., Adams,C., Bailey,M., Barberia,J., Blankenburg,K.,  
Bodette,B., Bouck,J., Bowler,S., Brooks,A., Buhay,C., Bunac,C.,  
Burkett,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,  
David,R., Delgado,O., Deshazo,D., Ding,Y., Domah-Rashid,N.,  
Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,  
Forcum-Taney,J., Frantz,P., Ganesh,R., Gorrell,J.H., Gorrell,L.L.,  
Guerara,W., Harris,K., Hernandez,J., Hodgson,A., Hogues,M.,  
Holloway,C., Hosak,H., Jackson,L.E., Jackson,L., Jia,Y., Jones,M.,  
Kelly,S., Kondejewski,N., Kong,Y., Kovar,C., Leal,B., Li,Z.,  
Lichterage,O., Liu,J., Liu,W., Logan,O., Lozado,R.U., Lu,U.,  
Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G., Morgan,M.,  
Morris,S., Nash,S., Nelson,A., Nguyen,R., Nguyen,S.,  
Oswal,G., Parish,B., Paxton,S., Payton,B., Perez,L., Pu,L.L.,  
Quiles,M., Reiter,D., Rives,M., Samuel,S., Say,J., Scherer,S.,  
Shah,B., Shen,H., Simon,M., Sparks,A., Stamps,A., Sugang,R.,  
Taber,P., Taylor,T., Vaequez,L., Vinson,R., Vo,Q., Wabahal,M.,  
Watlington,S., Weinstein,G., Weinstein,I.R., Williamson,A.,  
Worley,K., Wren,J., Wrenford,G., Yu,W., Zhou,X., Nelson,D. and  
Gibbs,R.  
Direct Submission  
Unpublished  
2 (bases 1 to 135805)  
Worley,K.C.  
TITLE Direct Submission  
AUTHORS Submitted (02-JUN-2000) Human Genome Sequencing Center, Department  
JOURNAL

## COMMENT

of Molecular and Human Genetics, Baylor College of Medicine, One  
Baylor Plaza, Houston, TX 77030, USA  
On Jul 25, 2000 this sequence version replaced gi:8468798.  
----- Genome Center  
Center: Baylor College of Medicine  
Center code: BCM  
Web site: <http://www.hgsc.bcm.tmc.edu/>  
Contact: [hgsc-help@bcm.tmc.edu](mailto:hgsc-help@bcm.tmc.edu)  
----- Project Information  
Center project name: HBR  
Center clone name: RP11-224122  
----- Summary Statistics  
Assembly program: Phrap; version 0.990329  
Consensus quality: 113598 bases at least Q40  
Consensus quality: 123530 bases at least Q20  
Consensus quality: 129169 bases at least Q20  
Estimated insert size: 126394; sum-of-contigs estimation  
Quality coverage: 0x in Q20 bases; agarose-fp estimation  
Quality coverage: 3.1x in Q20 bases; sum-of-contigs estimation  
-----  
\* NOTE: Estimated insert size may differ from sequence length  
\* (see [http://www.hgsc.bcm.tmc.edu/docs/genbank\\_draft\\_data.html](http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)).  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 26 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
1  
22050: contig of 22050 bp in length  
22150: gap of unknown length  
22151: contig of 12690 bp in length  
34840: gap of unknown length  
34841: contig of 11574 bp in length  
34941: gap of unknown length  
46515: gap of unknown length  
46516: contig of 14452 bp in length  
61067: gap of unknown length  
61167: contig of 7663 bp in length  
68849: gap of unknown length  
68850: gap of unknown length  
68949: gap of unknown length  
76117: contig of 7168 bp in length  
76217: gap of unknown length  
76218: contig of 7540 bp in length  
83758: gap of unknown length  
83857: gap of unknown length  
83858: contig of 5704 bp in length  
89561: gap of unknown length  
89562: gap of unknown length  
89661: gap of unknown length  
89662: contig of 6436 bp in length  
96097: gap of unknown length  
96098: gap of unknown length  
96197: gap of unknown length  
102149: contig of 5952 bp in length  
102150: gap of unknown length  
102249: gap of unknown length  
106020: contig of 3771 bp in length  
106120: gap of unknown length  
106121: contig of 4013 bp in length  
110133: gap of unknown length  
110134: contig of 3721 bp in length  
110234: gap of unknown length  
110235: contig of 3721 bp in length  
113955: gap of unknown length  
113956: contig of 2242 bp in length  
114055: gap of unknown length  
114056: gap of unknown length  
116296: gap of unknown length  
116297: gap of unknown length  
116397: contig of 1835 bp in length  
118331: gap of unknown length  
118332: gap of unknown length  
118333: contig of 1604 bp in length  
119935: gap of unknown length  
119936: gap of unknown length  
120035: gap of unknown length  
120036: contig of 1728 bp in length  
121764: gap of unknown length  
121765: gap of unknown length  
123764: contig of 1901 bp in length  
123765: gap of unknown length  
123865: gap of unknown length  
126122: contig of 2258 bp in length  
126123: gap of unknown length  
126222: gap of unknown length  
126223: contig of 1336 bp in length  
127558: gap of unknown length  
127559: gap of unknown length  
127658: contig of 1150 bp in length  
128808: gap of unknown length  
128908: gap of unknown length  
130088: contig of 1180 bp in length

\* 130089 130188: gap of unknown length  
 \* 130189 131521: contig of 1333 bp in length  
 \* 131522 131621: gap of unknown length  
 \* 131622 133147: contig of 1526 bp in length  
 \* 133148 133347: gap of unknown length  
 \* 133248 134356: contig of 1109 bp in length  
 \* 134357 134456: gap of unknown length  
 \* 134457 135805: contig of 1349 bp in length.  
 Location/Qualifiers  
 1. 135805

## FEATURES

source

/organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="3"  
 /clone="RP11-224122"  
 22051..22150  
 /estimated\_length=unknown  
 34841..34940  
 /estimated\_length=unknown  
 46515..46614  
 /estimated\_length=unknown  
 61067..61166  
 /estimated\_length=unknown  
 68850..68949  
 /estimated\_length=unknown  
 76118..76217  
 /estimated\_length=unknown  
 83758..83857  
 /estimated\_length=unknown  
 89562..89661  
 /estimated\_length=unknown  
 96098..96197  
 /estimated\_length=unknown  
 102150..102249  
 /estimated\_length=unknown  
 106021..106120  
 /estimated\_length=unknown  
 110134..110233  
 /estimated\_length=unknown  
 113955..114054  
 /estimated\_length=unknown  
 116297..116396  
 /estimated\_length=unknown  
 118232..118331  
 /estimated\_length=unknown  
 119336..120035  
 /estimated\_length=unknown  
 121764..121863  
 /estimated\_length=unknown  
 123765..123864  
 /estimated\_length=unknown  
 126123..126222  
 /estimated\_length=unknown  
 127559..127658  
 /estimated\_length=unknown  
 128809..128908  
 /estimated\_length=unknown  
 130089..130188  
 /estimated\_length=unknown  
 131522..131621  
 /estimated\_length=unknown  
 133148..133247  
 /estimated\_length=unknown  
 134357..134456  
 /estimated\_length=unknown

## ORIGIN

Query Match 1.7%; Score 52; DB 14; Length 135805;  
 Best Local Similarity 100.0%; Pred.No. 5.6e-16;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 2889 GAGGAGGTGATCACTGAGGCCAGAGTTGAGACCGAGCCTGGCCAACAT 2940  
 |||||||||||||||||||||||||||||||||||||||||||||||||||||

Db 85075 GAGGAGGTGATCACTGAGGCCAGAGTTGAGACCGAGCCTGGCCAACAT 85126

## RESULT 109

AC006006/C

## LOCUS

AC006006 136385 bp DNA linear PRI 27-JAN-2004  
 Homo sapiens PAC clone RP4-813P11 from 7, complete sequence.

## DEFINITION

AC006006

## ACCESSION

AC006006.2

## VERSION

GI:4753279

## KEYWORDS

HTG.

## SOURCE

Homo sapiens

## ORGANISM

Homo sapiens

## REFERENCE

AUTHORS

1 (bases 1 to 136385)

Hillier,L.W., Fulton,R.S., Fulton,L.A., Graves,T.A., Pepin,K.H.,

Wagner-McPherson,C., Layman,D., Maas,J., Jaeger,S., Walker,R.,

Wylie,K., Sekhon,M., Becker,M.C., O'Laughlin,M.D., Schaller,M.E.,

Du,H., Sun,H., Edwards,J., Bradshaw-Cordum,H., Ali,J., Andrews,S.,

Isak,A., Vandrun,A., Nguyen,C., Du,F., Lamar,B., Courtney,L.,

Kajicki,J., Ozersky,P., Bielicki,L., Scott,K., Holmes,A.,

Harkins,R., Harris,A., Strong,C.M., Hou,S., Tomlinson,C.,

Dauphin-Kohlberg,S., Kozlowski-Reilly,A., Leonard,S., Rohlfing,T.,

Rock,S.M., Tin-Wollam,A.M., Abbott,A., Minx,P., Maupin,R.,

Strommatt,C., Lathrelle,P., Miller,N., Johnson,D., Murray,J.,

Woessner,J.P., Wendt,M.C., Yang,S.P., Schultze,B.R., Wallis,J.W.,

Speleth,J., Bieri,T.A., Nelson,J.O., Berkowitz,N., Wohldmann,P.E.,

Cook,L.L., Hickenbotham,M.T., Eldred,J., Williams,D., Bedell,J.A.,

Mardis,E.R., Clifton,S.W., Chissee,S.L., Marra,M.A., Raymond,C.,

Haugen,E., Gillett,W., Zhou,Y., James,R., Phelps,K., Iadonoto,S.,

Bubb,K., Simms,E., Levy,R., Clendenning,J., Kaul,R., Kent,W.J.,

Pruy,J., Suyama,M., Bailey,J.A., Portnoy,M.E., Torrens,D.,

Chinwalla,A.T., Gish,W.R., Eddy,S.R., McPherson,J.D., Olson,M.V.,

Bichler,E.E., Green,E.D., Waterston,R.H. and Wilson,R.K.

The DNA sequence of human chromosome 7

Nature 424 (6945), 157-164 (2003)

12853948

2 (bases 1 to 136385)

Drone,K., Le T.P. and Wohldmann,P.

The sequence of Homo sapiens PAC clone RP4-813P11

Unpublished (2001)

3 (bases 1 to 136385)

Waterston,R.H.

Direct Submission

Submitted (22-NOV-1998) Genome Sequencing Center, Washington

University School of Medicine, 4444 Forest Park Parkway, St. Louis,

MO 63108, USA

4 (bases 1 to 136385)

Waterston,R.H.

Direct Submission

Submitted (05-MAY-1999) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 5 (bases 1 to 136385)  
 Waterston,R.  
 Direct Submission  
 Submitted (21-DEC-1999) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 6 (bases 1 to 136385)  
 Waterston,R.  
 Direct Submission  
 Submitted (27-JAN-2004) Department of Genetics, Washington  
 University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 On May 5, 1999 this sequence version replaced gi:3907515.  
 ----- Genome Center

Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu>  
Contact: [sapiens@wustl.wustl.edu](mailto:sapiens@wustl.wustl.edu)  
----- Summary Statistics -----  
Center project name: H\_DJ0813F11

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

#### MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu>

#### SOURCE INFORMATION:

This clone was derived from human PAC library RPCI-4, prepared by Pieter de Jong and coworkers at <http://www.chori.org> using the method described by Ioannou et al., Nature Genetics 6:84-9 (1994). The library is from one male donor.  
The clone may be obtained either from Genome Systems, Inc. (<http://www.genomesystems.com>) or Research Genetics, Inc. (<http://www.resgen.com>); or from Pieter de Jong.  
VECTOR: pCYPAC2

#### NEIGHBORING SEQUENCE INFORMATION:

The clone sequenced to the left is RP5-839B19, 200 bp overlap the clone sequenced to the right is RP5-1007M8. Actual start of this clone is at base position 197 of RP4-813F11 actual end is at 136385 of RP4-813F11.

#### FEATURES

##### source

1..136385  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="7"  
/map="7"  
/clone="RP4-813F11"  
/clone\_1lb="RPCI-4"  
1..91  
/rpt\_family="Alu"  
repeat\_region  
103..458  
/rpt\_family="MER2\_type"  
repeat\_region  
452..538  
/rpt\_family="MER2\_type"  
repeat\_region  
561..664  
/rpt\_family="Alu"  
repeat\_region  
665..1267  
/rpt\_family="MER2\_type"  
repeat\_region  
1342..1376  
/rpt\_family="Mariner"  
repeat\_region  
1494..1799  
/rpt\_family="Alu"  
repeat\_region  
1833..2172  
/rpt\_family="Alu"  
repeat\_region  
2948..3060  
/rpt\_family="MIR"

snRNA 3194..3300  
/product="snRNA-U6-related"  
repeat\_region 3194..3300  
/rpt\_family="U6"  
repeat\_region 3301..3345  
/rpt\_family="AT-rich"  
repeat\_region 3552..3598  
/rpt\_family="(TGG)n"  
repeat\_region 3611..3870  
/rpt\_family="Alu"  
repeat\_region 3871..4096  
/rpt\_family="MER1\_type"  
repeat\_region 4097..4424  
/rpt\_family="Alu"  
repeat\_region 4425..4657  
/rpt\_family="MER1\_type"  
repeat\_region 4660..4841  
/rpt\_family="L2"  
repeat\_region 4921..5043  
/rpt\_family="L2"  
repeat\_region 5310..5872  
/rpt\_family="L2"  
repeat\_region 6007..6151  
/rpt\_family="L2"  
repeat\_region 6843..7002  
/rpt\_family="Alu"  
repeat\_region 7003..7294  
/rpt\_family="Alu"  
repeat\_region 7307..7672  
/rpt\_family="MER1\_type"  
repeat\_region 7680..7971  
/rpt\_family="Alu"  
repeat\_region 7981..8122  
/rpt\_family="MER1\_type"  
repeat\_region 8154..8213  
/rpt\_family="MER2\_type"  
repeat\_region 8214..8274  
/rpt\_family="MER1\_type"  
repeat\_region 8333..8454  
/rpt\_family="MER1\_type"  
repeat\_region 8997..9108  
/rpt\_family="MIR"  
repeat\_region 9144..9340  
/rpt\_family="MIR"  
repeat\_region 9489..9600  
/rpt\_family="L2"  
repeat\_region 9589..9732  
/rpt\_family="CR1"  
repeat\_region 9735..9898  
/rpt\_family="MER1\_type"  
repeat\_region 10317..10376  
/rpt\_family="L2"  
repeat\_region 10406..10713  
/rpt\_family="MER2\_type"  
repeat\_region 10720..10879  
/rpt\_family="MER1\_type"  
repeat\_region 11045..11096  
/rpt\_family="MER103"  
repeat\_region 11097..11393

Query Match 1.7%; Score 52; DB 8; Length 136385;  
Best Local Similarity 100.0%; Pred. No. 5.6e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCAAGTGAATCACTGAGGCCAGGAGTTGAGACCAAGCTGGCCAACT 2940  
|||||  
DB 114432 GAGGCAAGTGAATCACTGAGGCCAGGAGTTGAGACCAAGCTGGCCAACT 114381

RESULT 110  
AC010367/c AC010367 137845 bp DNA 1linear HTG 20-APR-2001  
DEFINITION Homo sapiens chromosome 5 clone CTD-2044L23, WORKING DRAFT

SEQUENCE, 17 unordered pieces.  
ACCESSION AC010367  
VERSION AC010367.4 GI:13699509  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_ACTIVEPIN.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo  
1 (bases 1 to 137845)  
DOE Joint Genome Institute.  
Sequencing of Human Chromosome 5  
Unpublished  
2 (bases 1 to 137845)  
DOE Joint Genome Institute.  
Direct Submission  
Submitted (15-SEP-1999) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
On Apr 20, 2001 this sequence version replaced gi:7710742.  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: <http://www.jgi.doe.gov>  
-----  
Project Information  
Center Project Name: 644732  
Center clone name: CITB-H1\_2044L23  
-----  
Summary Statistics  
Consensus quality: 122756 bases at least Q40  
Consensus quality: 127174 bases at least Q30  
Consensus quality: 129642 bases at least Q20  
Estimated insert size: 121000; pulse field gel estimation  
Estimated insert size: 136245; sum-of-coverage estimation  
Quality coverage: 6.63 in Q20 bases; pulse field gel estimation  
Quality coverage: 5.89 in Q20 bases; sum-of-coverage estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 17 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\*  
\* 1 1041: contig of 1041 bp in length  
\* 1042 1141: gap of unknown length  
\* 1142 2183: contig of 1042 bp in length  
\* 2184 2283: gap of unknown length  
\* 2284 3529: contig of 1246 bp in length  
\* 3530 3629: gap of unknown length  
\* 3630 4897: contig of 1268 bp in length  
\* 4898 4997: gap of unknown length  
\* 4998 6540: contig of 1543 bp in length  
\* 6541 7735: gap of unknown length  
\* 7736 7835: contig of 1095 bp in length  
\* 7836 8852: gap of unknown length  
\* 8853 8952: contig of 1017 bp in length  
\* 8953 10065: gap of unknown length  
\* 10066 10165: contig of 1113 bp in length  
\* 10166 11335: gap of unknown length  
\* 11336 11435: contig of 1170 bp in length  
\* 11436 12444: gap of unknown length  
\* 12445 12544: contig of 1009 bp in length  
\* 12545 13852: gap of unknown length  
\* 13853 13952: contig of 1308 bp in length  
\* 13953 15056: gap of unknown length  
\* 15057 15156: contig of 1104 bp in length  
\* 15157 23003: gap of unknown length  
\* 23004 23103: contig of 7847 bp in length  
\* 23104 41386: gap of unknown length  
\* 41387 41486: contig of 18283 bp in length  
\* 41487 68942: gap of unknown length  
\* 68942 27456: contig of 27456 bp in length

FEATURES  
source \* 68943 69042: gap of unknown length  
\* 69043 98301: contig of 29259 bp in length  
\* 98302 98401: gap of unknown length  
\* 98402 137845: contig of 39444 bp in length.  
Location/Qualifiers  
1. 137845  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="CTD-2044L23"  
/clone\_1fb="Caltech human BAC library D"  
1042. 1141  
/estimated\_length=unknown  
2184. 2283  
/estimated\_length=unknown  
3530. 3629  
/estimated\_length=unknown  
4898. 4997  
/estimated\_length=unknown  
5541. 6640  
/estimated\_length=unknown  
7736. 7835  
/estimated\_length=unknown  
8853. 8952  
/estimated\_length=unknown  
10066. 10165  
/estimated\_length=unknown  
11336. 11435  
/estimated\_length=unknown  
12445. 12544  
/estimated\_length=unknown  
13853. 13952  
/estimated\_length=unknown  
15057. 15156  
/estimated\_length=unknown  
23004. 23103  
/estimated\_length=unknown  
41387. 41486  
/estimated\_length=unknown  
68943. 69042  
/estimated\_length=unknown  
98302. 98401  
/estimated\_length=unknown  
ORIGIN  
Query Match 1.7%; Score 52; DB 14; Length 137845;  
Best Local Similarity 100.0%; Pred.No. 5.6e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2899 GATCACCTGAGGCCGAGATTGAGACCGCTGCGCAACATAGCGAAACCC 2950  
|||||  
DB 33124 GATCACCTGAGGCCGAGATTGAGACCGCTGCGCAACATAGCGAAACCC 33073  
|||||  
RESULT 111  
AL954211 138573 bp DNA linear PRI 16-OCT-2003  
LOCUS AL954211  
DEFINITION Pan troglodytes chromosome 22 clone PFB-034605 map 22q22.11,  
complete sequence.  
ACCESSION AL954211  
VERSION AL954211.1 GI:37605794  
KEYWORDS HTG.  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Pan.  
REFERENCE 1 (bases 1 to 138573)  
\* The Chimpanzee Chromosome 22 Sequencing Consortium  
DNA sequence of chimpanzee chromosome 22 and its evolutionary  
implications



JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

unpublished  
2 (bases 1 to 138573)  
Kube,M., Sudbrak,R., Borzym,K., Mueller,I., Klages,S., Burgdorf,C.,  
Helmann,K., Gimel,V., Beck,A., Ben Kahla,A., Lehnach,H.,  
Yaspo,M.L. and Reinhardt,R.  
Direct Submission  
Submitted (18-DEC-2002) Max-Planck-Institute for Molecular  
Genetics, Department of Vertebrate Genomics, Ihnestrasse 73,  
Berlin, 14195 Germany  
The Chimpanzee Chromosome 22 Sequencing Consortium consists of:  
\*Chinese National Human Genome Center at Shanghai,  
Shanghai, China;  
\*GDF, Dept. of Genome Analysis, Braunschweig, Germany;  
\*Institute of Molecular Biotechnology, Jena, Germany;  
\*KIBB Genome Research Center, Daejeon, Korea;  
\*Max-Planck-Institute for Molecular Genetics, Berlin, Germany;  
\*National Institute of Genetics, Mishima, Japan;  
\*National Yang Ming University Genome Research Center, Taipei,  
Taiwan;  
\*RIKEN Genomic Sciences Center, Yokohama, Japan.

----- Genome Center  
Center: Max-Planck-Institute for Molecular Genetics  
Center code: MPIMG  
----- Project Information  
Center clone name: PTB-034G05  
----- Summary Statistics  
Sequencing vector: pUC18; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 13857 bases at least Q40  
Consensus quality: 138569 bases at least Q30  
Consensus quality: 138573 bases at least Q20  
Quality coverage: 13.72

-----  
This sequence was finished as follows unless otherwise noted: all  
regions were double stranded, sequenced with an alternate  
chemistry, or covered by high quality data (i.e., phred quality >=  
30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by at least  
one plasmid subclone or more than one M13 subclone; and the  
assembly was confirmed by restriction digest. -----  
Neighboring sequence information:  
This clone is overlapped by  
RP43-012J05  
RP43-043L10  
RP43-043L10

-----  
SOURCE INFORMATION:  
The PTB1 chimpanzee BAC library was prepared from DNA isolated from  
cultured cells established from the blood of a single male  
chimpanzee. Clones may be obtained from Asao Fujiyama and  
co-workers (http://www.gsc.riken.go.jp).  
VECTOR: pKS145.

-----  
FEATURES  
source  
1.138573  
Location/Qualifiers  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/map="22q22.11"  
/chromosome="22"  
/clone="PTB-034G05"  
-----  
1.12748  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone="RP43-012J05"  
/note="overlapping clone"  
-----  
96436..>138573  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone="RP43-043L10"  
/note="overlapping clone"

source

ORIGIN

Query Match 1.7%; Score 52; DB 8; Length 138573;  
Best Local Similarity 100.0%; Pred. No. 5.6e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3071 CAAGATTGTGCACCTGCAGCTCGGCAACAGACCAAGACTGTCTC 3122  
|||||  
112046 CAAGATTGTGCACCTGCAGCTCGGCAACAGACCAAGACTGTCTC 112097  
|||||

RESULT 112  
AC073542 141633 bp DNA linear PRI 01-MAY-2001  
LOCUS Homo sapiens chromosome 5 clone RP11-221L16, complete sequence.  
DEFINITION AC073542  
ACCESSION AC073542.4 GI:13876495  
VERSION HTG.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 141633)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
JOURNAL Unpublished  
AUTHORS DOE Joint Genome Institute.  
REFERENCE 2 (bases 1 to 141633)  
DIRECT SUBMISSION  
Submitted (22-JUN-2000) Production Sequencing Facility, DOE Joint  
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 141633)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
AUTHORS Direct Submission  
Submitted (28-APR-2001) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
4 (bases 1 to 141633)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
JOURNAL Direct Submission  
Submitted (01-MAY-2001) DOE Joint Genome Institute, 2800 Mitchell  
Drive, Walnut Creek, CA 94598, USA  
On Apr 28, 2001 this sequence version replaced gi:13677059.  
COMMENT Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www.shgc.stanford.edu  
Quality: Phrap Quality >=40 99.9% of Sequence;  
Estimated Total Number of Errors is 0.1.

-----  
FEATURES  
source  
1.141633  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="RP11-221L16"

ORIGIN

Query Match 1.7%; Score 52; DB 8; Length 141633;  
Best Local Similarity 100.0%; Pred. No. 5.6e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3071 CAAGATTGTGCACCTGCAGCTCGGCAACAGACCAAGACTGTCTC 3122  
|||||  
105940 CAAGATTGTGCACCTGCAGCTCGGCAACAGACCAAGACTGTCTC 105991  
|||||

RESULT 113  
AL357552 143969 bp DNA linear PRI 18-MAY-2005  
LOCUS Human DNA sequence from clone RP4-63318 on chromosome 1p36.21-36.32  
DEFINITION Contains the 5' end of the RERB gene for arginine-glutamic acid  
dipeptide (RE) repeats (F1438775), two ribosomal protein L7 (RPL7)  
pseudogenes, a ribosomal protein L27 (RPL27) pseudogene, a novel

ACCESSION AL357552 GI:27652622  
 VERSION AL357552.25  
 KEYWORDS HTG; FLJ38775; RERR; RPL27; RPL7.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo  
 1 (bases) to 143969  
 Davidson, C.  
 Direct Submission  
 Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk  
 Clone requests: clonerequest@sanger.ac.uk  
 On Jan 11, 2003 this sequence version replaced gi:27498402.  
 The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:  
 Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at  
 http://www.sanger.ac.uk/projects/C.elegans/wormpep  
 This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at  
 http://www.sanger.ac.uk/HGP/Chr1  
 RP4-63318 is from the library RPCI-4 constructed by the group of Plieter de Jong. For further details see  
 http://www.choi.org/bacpac/home.htm  
 VECTOR: pCYPAC2  
 ----- Genome Center  
 Center: Wellcome Trust Sanger Institute  
 Center code: SC  
 Web site: http://www.sanger.ac.uk  
 Contact: vegas@sanger.ac.uk  
 -----  
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.  
 Location/Qualifiers  
 1..143969  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="R2PD:RPCIP704108633"  
 /db\_xref="taxon:9606"  
 /chromosome="1"  
 /map="p36.21-36.32"  
 /clone="RP4-63318"  
 /clone\_11b="RPCI-4"  
 1  
 /note="Clone\_left\_end: RP4-63318"  
 misc\_feature  
 8344  
 /note="Clone\_right\_end: RPL1-3323"  
 join(complement(16843..17007),  
 complement(AL357713.11:73394..73642))  
 /gene="RERR"  
 locus\_tag="RPL1-141M15.2-009"  
 join(complement(16843..17007),  
 complement(AL357713.11:73394..73642))  
 /gene="RERR"  
 locus\_tag="RPL1-141M15.2-009"  
 /product="arginine-glutamic acid dipeptide (RE) repeats"  
 /note="match: ESTs; Em:A1524637.1 Em:AW444820.1 Em:BB675021.1 Em:BF512881.1"  
 64218..64875  
 /locus\_tag="RP4-63318.1-001"  
 /pseudo  
 64218..64875  
 /locus\_tag="RP4-63318.1-001"  
 /note="match: proteins; SW:P05426 SW:P14148 SW:P18124

Tr:AAH25909 Tr:BAC34366 Tr:BAC40262 Tr:O95036 Tr:Q15289  
 Tr:O80UR0 Tr:O90YW3 Tr:Q9D6G9"  
 /pseudo  
 /codon\_start=1  
 /product="ribosomal protein L7 (RPL7) pseudogene"  
 join(complement(67219..67626),  
 complement(AL357713.11:73174..73642),  
 complement(AL357713.11:41511..41581),  
 complement(AL357713.11:31762..31887),  
 complement(AL357713.11:152922..153027),  
 complement(AL357713.11:151979..152075),  
 complement(AL357713.11:136718..136822),  
 complement(AL357713.11:104131..104179),  
 complement(AL357713.11:93034..93044),  
 complement(AL357713.11:90568..90667),  
 complement(AL357713.11:60475..61528))  
 /gene="RERR"  
 locus\_tag="RPL1-141M15.2-010"  
 join(complement(67219..67626),  
 complement(AL357713.11:73174..73642),  
 complement(AL357713.11:41511..41581),  
 complement(AL357713.11:31762..31887),  
 complement(AL357713.11:152922..153027),  
 complement(AL357713.11:151979..152075),  
 complement(AL357713.11:136718..136822),  
 complement(AL357713.11:104131..104179),  
 complement(AL357713.11:93034..93044),  
 complement(AL357713.11:90568..90667),  
 complement(AL357713.11:60475..61528))  
 /gene="RERR"  
 locus\_tag="RPL1-141M15.2-010"  
 /product="arginine-glutamic acid dipeptide (RE) repeats"  
 /note="match: CDNs; Em:AK06094.1"  
 87829..87916  
 /note="Single clone region. Assembly confirmed by restriction digest data."  
 94628  
 /note="Clone\_left\_end: RPL1-185B14"  
 96733..96734  
 /note="Sequence from overlapping clone RPL1-185B14 (AL354930). Assembly confirmed by restriction digest."  
 complement(99991..100723)  
 locus\_tag="RP4-63318.2-001"  
 /pseudo  
 complement(99991..100723)  
 locus\_tag="RP4-63318.2-001"  
 /product="ribosomal protein L7 (RPL7) pseudogene"\n"  
 join(complement(130948..131253),  
 complement(106192..106376),  
 complement(AL357713.11:73174..73642),  
 complement(AL357713.11:41511..41581),  
 complement(AL357713.11:31762..31887),  
 complement(AL357713.11:152922..153027),  
 complement(AL357713.11:151979..152075),  
 complement(AL357713.11:136718..136822),  
 complement(AL357713.11:104131..104179),  
 complement(AL357713.11:93034..93044),  
 complement(AL357713.11:90568..90667),  
 complement(AL357713.11:60475..61528),  
 complement(AL357713.11:18232..18312),  
 complement(AL357713.11:80582..80649),  
 complement(AL357713.11:79899..80091),  
 complement(AL357713.11:78600..78660),  
 complement(AL357713.11:77599..77726),  
 complement(AL357713.11:75946..77326),  
 complement(AL357713.11:75600..75822),  
 complement(AL357713.11:74033..74752),  
 complement(AL357713.11:71936..72082),

complement (AL356072.21:71255..71435)  
 complement (AL356072.21:68233..70956)  
 /gene="RERE"  
 /locus\_tag="RP11-141M15.2-001"  
 join(complement(130948..131253),  
 complement(106192..106376),  
 complement (AL357713.11:71174..73642),  
 complement (AL357713.11:41511..41541),  
 complement (AL357713.11:31762..31887),  
 complement (AL096855.48:152922..153027),  
 complement (AL096855.48:151979..152075),  
 complement (AL096855.48:136718..136822),  
 complement (AL096855.48:104131..104179),  
 complement (AL096855.48:92910..93034),  
 complement (AL096855.48:90568..90657),  
 complement (AL096855.48:61430..61528),  
 complement (AL096855.48:18232..18312),  
 complement (AL356072.21:81648..81810),  
 complement (AL356072.21:80582..80674),  
 complement (AL356072.21:79892..80091),  
 complement (AL356072.21:78519..78680),  
 complement (AL356072.21:77599..77712),  
 complement (AL356072.21:75948..77326),  
 complement (AL356072.21:75600..75822),  
 complement (AL356072.21:74032..74752),  
 complement (AL356072.21:71936..72082),  
 complement (AL356072.21:71255..71435),  
 complement (AL356072.21:68233..70956)  
 /gene="RERE"  
 /locus\_tag="RP11-141M15.2-001"  
 /product="arginine-glutamic acid dipeptide (RE) repeats"  
 /note="match: CDNA5: Em:AB036737.1"  
 join(112263..112479,113097..113294)  
 /locus\_tag="RP4-63318.3-001"  
 /pseudo  
 join(112263..112479,113097..113294)  
 /locus\_tag="RP4-63318.3-001"  
 /note="match: proteins: SW:PO8526 Tr:AAH24366 Tr:BAH28321  
 Tr:Q8WY13 Tr:Q80YU1 Tr:Q9DFL2"  
 /pseudo  
 /codon\_start=1  
 /product="ribosomal protein L27 (RPL27) pseudogene"  
 115538..115554  
 /note="Sequence from overlapping clone RP11-185B14  
 (AL354930) Assembly confirmed by restriction digest."  
 join(119648..119726,120566..120839)  
 /locus\_tag="RP4-63318.4-001"  
 join(119648..119726,120566..120839)  
 /locus\_tag="RP4-63318.4-001"  
 /locus\_tag="RP4-63318.4-001"  
 Query Match 1.7% Score 52; DB 8; Length 143969;  
 Best Local Similarity 100.0%; Pred. No. 5.6e-16;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 3071 CAGATTGTCACCTGCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
 |||||  
 Db 50064 CAAAGTTGGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 50013  
 |||||  
 RESULT 114  
 AC027146 AC027146 144542 bp DNA 1linear HTG 28-AUG-2002  
 LOCUS Homo sapiens chromosome 17 clone RP11-123N2 map 17, WORKING DRAFT  
 DEFINITION  
 SEQUENCE, 28 unordered pieces.  
 AC027146 AC027146  
 VERSION AC027146.1 GI:7331426  
 KEYWORDS HTG, HTGS PHASE1, HTGS\_DRAFT.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 144542)

AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
 Homo sapiens chromosome 17, clone RP11-123N2  
 Unpublished  
 2 (bases 1 to 144542)  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedalov, F.,  
 Boguslavsky, L., Boudgaltier, B., Brown, A., Burkett, G.,  
 Campiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,  
 Collymore, A., Cooke, P., Dekrelano, K., Dewar, K., Diaz, J. S.,  
 Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
 Gargan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
 Grand, P., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
 Howard, J. C., Iliev, I., Johnson, R., Jones, C., Kam, L., Karakas, A.,  
 Klein, J., Labouque, K., Lamazares, R., Landers, T., Lechoczy, J.,  
 Levine, R., Lieu, C., Liu, G., Locke, K., McDonald, P., Margulis, N.,  
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R.,  
 Melidrim, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,  
 Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,  
 O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,  
 Pisanu, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talmas, J.,  
 Testa, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, U.,  
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,  
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (28-MAR-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 144542)  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bedalov, F.,  
 Boguslavsky, L., Boudgaltier, B., Brown, A., Burkett, G.,  
 Campiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,  
 Collymore, A., Cooke, P., Dekrelano, K., Dewar, K., Diaz, J. S.,  
 Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
 Gargan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
 Grand, P., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
 Howard, J. C., Iliev, I., Johnson, R., Jones, C., Kam, L., Karakas, A.,  
 Klein, J., Labouque, K., Lamazares, R., Landers, T., Lechoczy, J.,  
 Levine, R., Lieu, C., Liu, G., Locke, K., McDonald, P., Margulis, N.,  
 McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheters, R.,  
 Melidrim, J., Meneus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,  
 Murphy, T., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,  
 O'Neill, D., Oliver, T. M., Oliver, J., Peterson, K., Pierre, N.,  
 Pisanu, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talmas, J.,  
 Testa, S., Theodore, J., Tirrell, A., Travers, M., Triggillo, U.,  
 Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J.,  
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.  
 Direct Submission  
 Submitted (28-AUG-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 COMMENT  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L5016  
 Center clone name: 123 N 2  
 Summary Statistics  
 Sequencing vector: M13; W77815; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 128149 bases at least Q40  
 Consensus quality: 136286 bases at least Q30  
 Consensus quality: 139579 bases at least Q20  
 Insert size: 168000; agarose-fp  
 Insert size: 141842; sum-of-contigs

Quality coverage: 3.5 in Q20 bases; agarose-fp  
Quality coverage: 4.2 in Q20 bases; sum-of-coverage

NOTE: This is a 'working draft' sequence. It currently consists of 28 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```

1 1042: contig of 1042 bp in length
1043 1142: gap of 100 bp
1143 2215: contig of 1073 bp in length
2216 2316: gap of 100 bp
2316 3520: contig of 1205 bp in length
3521 3620: gap of 100 bp
3621 6235: contig of 2615 bp in length
6236 8261: contig of 1926 bp in length
8262 8361: gap of 100 bp
8362 10203: contig of 1842 bp in length
10204 10303: gap of 100 bp
10304 11741: contig of 1438 bp in length
11742 11841: gap of 100 bp
11842 13740: contig of 1899 bp in length
13741 13840: gap of 100 bp
13841 16452: contig of 2612 bp in length
16453 16552: gap of 100 bp
16553 19320: contig of 2768 bp in length
19321 19420: gap of 100 bp
19421 21340: contig of 1920 bp in length
21341 21440: gap of 100 bp
21441 23790: contig of 2350 bp in length
23791 23890: gap of 100 bp
23891 25590: contig of 1700 bp in length
25591 25690: gap of 100 bp
25691 28404: contig of 2714 bp in length
28405 28504: gap of 100 bp
28505 30466: contig of 1962 bp in length
30467 30566: gap of 100 bp
30567 34336: contig of 3770 bp in length
34337 34436: gap of 100 bp
34437 37626: contig of 3190 bp in length
37627 37726: gap of 100 bp
37727 41203: contig of 3477 bp in length
41204 41303: gap of 100 bp
41304 44513: contig of 3210 bp in length
44514 44613: gap of 100 bp
44614 48707: contig of 4094 bp in length
48708 48807: gap of 100 bp
48808 54126: contig of 5319 bp in length
54127 54226: gap of 100 bp
54227 62040: contig of 7814 bp in length
62041 62140: gap of 100 bp
62141 71155: contig of 9015 bp in length
71156 71255: gap of 100 bp
71256 80764: contig of 9509 bp in length
80765 80864: gap of 100 bp
80865 91109: contig of 10245 bp in length
91110 91209: gap of 100 bp
91210 101437: contig of 10228 bp in length
101438 101537: gap of 100 bp
101538 121402: contig of 19865 bp in length
121403 121502: gap of 100 bp
121503 144542: contig of 23040 bp in length.
Location/Qualifiers
1..144542
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="17"
/map="17"

```

```

/misc_feature /clone="RP11-123N2"
gap /clone_1b="RP11-11 Human Male BAC"
1..1042
/misc_feature /note="assembly_fragment"
gap 1043..1142
/misc_feature /estimated_length=100
gap 1143..2215
/misc_feature /note="assembly_fragment"
gap 2216..2315
/misc_feature /estimated_length=100
gap 2316..3520
/misc_feature /note="assembly_fragment"
gap 3521..3620
/misc_feature /estimated_length=100
gap 3621..6235
/misc_feature /note="assembly_fragment"
gap 6236..8261
/misc_feature /note="assembly_fragment"
gap 8262..8361
/misc_feature /estimated_length=100
gap 8362..10203
/misc_feature /note="assembly_fragment"
gap 10204..10303
/misc_feature /estimated_length=100
gap 10304..11741
/misc_feature /note="assembly_fragment"
gap 11742..11841
/misc_feature /estimated_length=100
gap 11842..13740
/misc_feature /note="assembly_fragment"
gap 13741..13840
/misc_feature /estimated_length=100
gap 13841..16452
/misc_feature /note="assembly_fragment"
gap 16453..16552
/misc_feature /estimated_length=100
gap 16553..19320
/misc_feature /note="assembly_fragment"
gap 19321..19420
gap /estimated_length=100

```

Query Match 1.7% Score 52; DB 14; Length 144542;  
Best Local Similarity 100.0%; Pred. No. 5.6e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 3071 CAAGATTGTGCACTGCACTCCAGGCTGGGCAAGACAGACAGCTGTCTC 3122
DB 110687 CAAGATTGTGCACTGCACTCCAGGCTGGGCAAGACAGACAGCTGTCTC 110738

```

#### RESULT 115

AC110281 146199 bp DNA linear HTG 27-MAR-2003  
LOCUS Homo sapiens chromosome 11 clone CTD-2132H1 map 11, 6 ordered  
DEFINITION

#### AC110281

#### ACCESSION

#### VERSION

#### KEYWORDS

#### SOURCE

#### ORGANISM

#### REFERENCE

#### AUTHORS

#### TITLE

#### JOURNAL

#### REFERENCE

#### AUTHORS

#### TITLE

#### JOURNAL

#### REFERENCE

1 (bases 1 to 146199)  
Britten, B., Linton, L., Nussbaum, C., Lander, E., All, A., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B.,  
Brown, A., Camarata, D., Campopiano, A., Chang, J., Chazaro, B.,  
Hominidae; Homo.  
Mammalia; Butcheria; Euarctontoglires; Primates; Catarrhini;  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Hominidae; Homo.  
1 (bases 1 to 146199)  
Britten, B., Linton, L., Nussbaum, C., Lander, E., All, A., Allen, N.,  
Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B.,  
Brown, A., Camarata, D., Campopiano, A., Chang, J., Chazaro, B.,



of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
3 lbaaes 1 to 146376)  
Worley,K.C.  
Direct Submision  
Submitted (08-DEC-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
4 lbaaes 1 to 146376)  
Worley,K.C.  
Direct Submision  
Submitted (30-MAR-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
5 lbaaes 1 to 146376)  
Worley,K.C.  
Direct Submision  
Submitted (14-DEC-2001) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

COMMENT  
On Dec 14, 2001 this sequence version replaced gi:13489132.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

ANNOTATION OF FEATURES:  
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.  
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.  
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least 2 exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

QUALSTAT-REPORT.  
Location/Qualifiers  
1. 146376  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="RP11-410F19"  
789. .825  
repeat\_region  
/rpt\_family="AT\_rich"  
1371. .1711  
/note="Region: Similar to Hs#S1074034 0z46g11.x1 Homo sapiens cDNA, A1079997"  
1737. .1781  
repeat\_region  
/rpt\_family="CA)n"  
3500. .3520  
repeat\_region

misc\_feature  
/rpt\_family="AT\_rich"  
join(14659..4721,6288..6513)  
/note="Region: Similar to Hs#S1609703 DKFZp762L097\_r1 Homo sapiens cDNA, A1121466"  
5272. .5325  
repeat\_region  
/rpt\_family="GC\_rich"  
6178. .6278  
repeat\_region  
/rpt\_family="GC\_rich"  
6495. .6637  
repeat\_region  
/rpt\_family="C-rich"  
6695. .6718  
repeat\_region  
/rpt\_family="GC\_rich"  
7005. .7142  
repeat\_region  
/rpt\_family="G-rich"  
8361. .8487  
repeat\_region  
/rpt\_family="MIR"  
9046. .9067  
repeat\_region  
/rpt\_family="TTG)n"  
complement(9153..9513)  
/rpt\_family="TTR1B"  
complement(10270..10517)  
/rpt\_family="AluUc"  
complement(10736..11062)  
/rpt\_family="MER1B"  
1101. .1155  
repeat\_region  
/rpt\_family="MIR"  
11367. .11452  
repeat\_region  
/rpt\_family="L1MB8"  
complement(11696..11997)  
/rpt\_family="L2"  
12207. .12827  
repeat\_region  
/rpt\_family="HAL1"  
12980. .13267  
repeat\_region  
/rpt\_family="AluUc"  
13473. .13671  
repeat\_region  
/rpt\_family="L2"  
complement(14040..14312)  
/rpt\_family="AluUc"  
complement(14495..14807)  
/rpt\_family="AluUsp"  
complement(14897..15005)  
/rpt\_family="MER5A"  
complement(15090..15172)  
/rpt\_family="L1MD3"  
complement(15182..15221)  
/rpt\_family="L1MC4"  
15222. .15533  
repeat\_region  
/rpt\_family="AluUc"  
complement(15534..15650)  
/rpt\_family="L1MC4"  
15651. .15954  
repeat\_region  
/rpt\_family="AluUc5"  
complement(15955..16485)  
/rpt\_family="L1MC4"  
complement(16586..16678)  
/rpt\_family="AluUc/x"  
complement(16679..16966)  
/rpt\_family="L1MC4"  
17043. .17351  
repeat\_region  
/rpt\_family="AluUg1"  
17320. .17410  
misc\_feature  
/function="low coverage"  
17376. .17407  
repeat\_region  
/rpt\_family="TTTGG)n"  
complement(17408..17699)  
/rpt\_family="AluUc"  
complement(17700..17843)  
/rpt\_family="L1MC4"  
complement(17878..18184)  
/rpt\_family="AluUg"  
complement(18422..18733)  
/rpt\_family="AluUc"  
complement(18919..20097)  
repeat\_region



```

repeat_region      /rpt_family="L1ME"
                    20107..20275
repeat_region      /rpt_family="L1PA13"
                    20276..20580
repeat_region      /rpt_family="AluSx"
                    20581..20741
repeat_region      /rpt_family="L1PA13"
                    20756..20865
repeat_region      /rpt_family="HAL1"
                    20879..20913
repeat_region      /rpt_family="(TTTTA)n"
                    complement(20914..21210)
repeat_region      /rpt_family="AluSg"
                    complement(21253..21311)
repeat_region      /rpt_family="LTR33"
                    complement(21334..21422)
repeat_region      /rpt_family="MER2"

```

```

Query Match      1.7%: Score 52; DB 8; Length 146376;
Best Local Similarity 100.0%; Pred No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy      3071 CAAGATTGTGCACGTGCACTCCAGCTGGCAACAGAGCAAGACTCTGTCTC 3122
Db      37189 CAAGATTGTGCACGTGCACTCCAGCTGGCAACAGAGCAAGACTCTGTCTC 37240

```

```

RESULT 117
CNS01DIX      146740 bp      DNA      linear      PRI 28-JUN-2001
LOCUS      CNS01DIX/c
DEFINITION      Human chromosome 14 DNA sequence BAC C-2002H8 of library CalTech-D
ACCESSION      AL132989.5 GI:14586417
VERSION      HTG; HTGS; ACTIVEFIN.
KEYWORDS      Homo sapiens (human)
SOURCE      Homo sapiens
ORGANISM      Homo sapiens

```

```

REFERENCE      Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
AUTHORS      Homnidae; Homo.

```

```

1 (bases 1 to 146740)
Hellig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
Brottier,P., Catolico,L., Barbe,V., Pelletier,B., Artiguenave,F.,
Levy,M., Beckenbery,R., Brula,T., deBertandis,V., Cruaud,C.,
Gyapay,G., Saurin,W. and Weissenbach,J.
Sequencing of the human chromosome 14
Unpublished
2 (bases 1 to 146740)
Genoscope.

```

```

Direct Submission
Submitted (28-JUN-2001) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : secrete@genoscope.cns.fr)
Web : www.genoscope.cns.fr
On Jul 2, 2001 this sequence version replaced gi:12580632.
----- Genome Center
Center: Genoscope / Centre National de Sequencage
Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: secrete@genoscope.cns.fr

```

```

The following BAC sequence is oriented from the T7 to the SP6 end.
Upstream BAC (overlapping the T7 end) : R-349A22
Downstream BAC (overlapping the SP6 end) : R-517013 (AC=AL13021)
----- Summary Statistics
Assembly program: Phrap; version 2.0
Quality coverage: 8.73x in Q20 bases; sum-of-contigs
-----
Overall quality chart :
Range : bases
1 - 9 :
10 - 19 : 1

```

```

20 - 29 : 2
30 - 39 : 41
40 - 49 : 1267
50 - 59 : 3807
60 - 69 : 6119
70 - 79 : 18019
80 - 89 : 48401
90 - 99 : 69083
-----
Percentage of bases with a quality value >= 40 : 99 %.
Location/Qualifiers
1..146740
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="14"
/clone="C-2002H8"
/clone_1b="CalTech-D"
30201..30314
/note="matching EMBL:G20444
Rdb:RH17151
Rdb:RH16912
dbSTS:STS11150
Identified using the e-PCR software (G. Schuler)"
30335..30431
/note="matching EMBL:Z38484
Rdb:RH1710
Rdb:RH1724
dbSTS:STS64724
Identified using the e-PCR software (G. Schuler)"
61096..61338
/note="matching EMBL:N22004
Rdb:RH69274
dbSTS:STS49152
Identified using the e-PCR software (G. Schuler)"
61230..61323
/note="matching EMBL:G38638
Rdb:RH97264
Rdb:RH83920
dbSTS:STS59275
Identified using the e-PCR software (G. Schuler)"
63152..63253
/note="matching EMBL:R85973
Rdb:RH53753
Rdb:RH14480
dbSTS:STS1946
Identified using the e-PCR software (G. Schuler)"
67540..67664
/note="matching EMBL:T63600
Rdb:RH53926
Rdb:RH14889
dbSTS:STS8565
Identified using the e-PCR software (G. Schuler)"
67545..67664
/note="matching EMBL:G28318
Rdb:RH33309
dbSTS:STS8564
Identified using the e-PCR software (G. Schuler)"
74562..74629
/note="matching EMBL:H95046
Rdb:RH6861
dbSTS:STS48768
Identified using the e-PCR software (G. Schuler)"
96307..96430
/note="matching EMBL:G30132
Rdb:RH37908
dbSTS:STS22657
Identified using the e-PCR software (G. Schuler)"
134360..134499
/note="matching EMBL:G33047
Rdb:RH67742
dbSTS:STS47669
Identified using the e-PCR software (G. Schuler)"

```

ORIGIN



exon complement(19383. .19456)  
/gene="PM-scl"  
/number=20  
complement(19457. .20569)  
/gene="PM-scl"  
/number=19  
complement(20570. .20654)  
/gene="PM-scl"  
/number=19  
complement(20655. .22714)  
/gene="PM-scl"  
/number=18  
complement(22715. .22810)  
/gene="PM-scl"  
/number=18  
complement(22811. .25328)  
/gene="PM-scl"  
/number=17  
complement(25329. .25435)  
/gene="PM-scl"  
/number=17  
complement(25436. .25851)  
/gene="PM-scl"  
/number=16  
complement(25852. .25930)  
/gene="PM-scl"  
/number=16  
complement(25931. .26087)  
/gene="PM-scl"  
/number=15  
complement(26088. .26138)  
/gene="PM-scl"  
/number=15  
complement(26139. .28197)  
/gene="PM-scl"  
/number=14  
complement(28198. .28309)  
/gene="PM-scl"  
/number=14  
complement(28310. .28985)  
/gene="PM-scl"  
/number=13  
complement(28986. .29036)  
/gene="PM-scl"  
/number=13  
complement(29037. .29248)  
/gene="PM-scl"  
/number=12  
complement(29249. .29397)  
/gene="PM-scl"  
/number=12  
complement(29398. .29566)  
/gene="PM-scl"  
/number=11  
complement(29567. .29723)  
/gene="PM-scl"  
/number=11  
complement(29724. .31172)  
/gene="PM-scl"  
/number=10  
complement(31173. .31363)  
/gene="PM-scl"  
/number=10  
complement(31364. .35932)  
/gene="PM-scl"  
/number=9

Query Match 1.7% Score 52; DB 8; Length 148049;  
Best Local Similarity 100.0%; Pred.No.5.7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3071 CAGAGTTGTGCACTGCACCTCCAGCTGTGGCAACAGAGCAAGACTCTGTCTC 3122  
|||||

DB 70758 CAGAGTTGTGCACTGCACCTCCAGCTGTGGCAACAGAGCAAGACTCTGTCTC 70707

RESULT 119  
AC002351/C  
LOCUS  
DEFINITION Homo sapiens 12 PAC RPI-46F2 (Roswell Park Cancer Institute Human  
PAC library) complete sequence.  
ACCESSION AC002351.2 GI:28460712  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 149679)  
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,  
Alshrocks,S.L., Amarantunga,H.C., Are,J.R., Ayala,M., Banks,T.,  
Barbata,J., Benton,J., Blmage,K., Blankenburg,K., Bonnin,D.,  
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,  
Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,  
Carroll,T.F., Carter,M., Cavazos,S.R., Chacko,U., Chavez,D.,  
Chen,G., Chen,R., Chen,Z., Chiu,D., Chowdhry,I., Christopoulos,C.,  
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,  
Dayla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,  
Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,  
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,  
Earhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Emerling,S.,  
Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P.,  
Frantz,P., Gabist,A., Gao,J., Garcia,A., Garner,T., Garza,N.,  
Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S.,  
Hamilton,K., Han,J., Harris,C., Harris,K., Hart,M., Havak,P.,  
Hawes,A., Hernandez,J., Hernandez,O., Hodgson,A., Hogue,M.,  
Holloway,C., Hollins,B., Homel,F., Howard,S., Huber,J., Hulyk,S.,  
Hume,J., Ioshikhes,I., Jackson,L.E., Jacobson,B., Jia,Y.,  
Johnson,R., Jolivet,S., Joudah,S., Karlsson,B., Kelly,S., Khan,U.,  
King,L., Korah,J., Kovar,C., Kratochvil,J., Kureshi,A., Landry,N.,  
Leal,B., Lee,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O.,  
Lieu,C., Liu,J., Liu,W., Louisedge,H., Lozdo,R.U., Lu,X.,  
Luchter,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapa,P.,  
Marcondel,I., Martin,R., Martindale,A., Martinez,E., Massey,E.,  
Mawhinney,E., McLeod,M.P., Meador,M., Mei,G., Merscher,S.,  
Metzker,M., Miller,A., Miner,G., Miner,Z., Mitchell,T.,  
Mohabbat,K., Montgomery,K.T., Morgan,M., Morris,S., Moser,M.,  
Neal,D., Nelson,D., Newton,U., Newton,N., Nguyen,A., Nguyen,N.,  
Nguyen,N., Nickerson,E., Nwokkenwo,S., Ogun,M., Okwuonu,G.,  
Oregunye,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,  
Peters,L., Pickens,R., Primus,B., Pu,L.L., Quiles,M., Ren,Y.,  
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savary,G.,  
Scherer,S., Scott,G., Shen,H., Shm.C., Shooshbari,N., Sisson,I.,  
Sodergren,B., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,  
Sutton,A., Syatek,A., Tabor,P., Tameisla,A., Tameisla,K., Tang,H.,  
Tansley,J., Taylor,C., Taylor,T., Teltord,B., Thomas,S.,  
Uemami,K., Vasquez,L., Vera,V., Villalón,D., Vinson,R., Wang,Q.,  
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,  
Williams,G., Williamson,A., Wiczyski,R., Wooden,S., Worley,K.,  
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Kucherlapati,R.,  
Weinstock,G. and Gibbs,R.

TITLE  
JOURNAL  
AUTHORS  
REFERENCE  
JOURNAL  
TITLES  
JOURNAL  
AUTHORS  
REFERENCE

Unpublished  
Direct Submission  
2 (bases 1 to 149679)  
Chiu,M.W.  
Direct Submission  
Submitted (18-JUL-1997) Molecular and Human Genetics, Baylor  
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
3 (bases 1 to 149679)  
Worley,K.C.  
Direct Submission  
Submitted (27-JAN-1998) Molecular and Human Genetics, Baylor  
College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
4 (bases 1 to 149679)  
Worley,K.C.

TITLE Direct Submission  
JOURNAL Submitted (01-MAY-1998) Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
REFERENCE 5 (bases 1 to 149679)  
AUTHORS Worley, K.C.  
JOURNAL Direct Submission  
SUBMITTED (05-MAY-1998) Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
REFERENCE 6 (bases 1 to 149679)  
AUTHORS Worley, K.C.  
JOURNAL Direct Submission  
SUBMITTED (09-JUN-1998) Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
REFERENCE 7 (bases 1 to 149679)  
AUTHORS Worley, K.C.  
JOURNAL Direct Submission  
SUBMITTED (28-JUL-1998) Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
REFERENCE 8 (bases 1 to 149679)  
AUTHORS Worley, K.C.  
JOURNAL Direct Submission  
SUBMITTED (21-FEB-2003) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA  
COMMENT On Feb 21, 2003 this sequence version replaced gi:2811094.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the features listing.

ANNOTATION OF FEATURES:  
STSs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.  
Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.  
Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

FEATURES  
source Location/Qualifiers  
1. 149679  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="12"  
/clone="RP1-46P2"  
misc\_feature 1..220  
/note="overlaps bases 102372..102591 of clone AC002978"  
/function="clone overlap"  
36..62  
repeat\_region

/rpt\_family="(TCTG)n"  
763..879  
repeat\_region  
/rpt\_family="CT-rich"  
1019..1056  
repeat\_region  
/rpt\_family="(CAT)n"  
complement(1124..1203)  
repeat\_region  
/rpt\_family="MIR"  
complement(1558..1748)  
repeat\_region  
/rpt\_family="MIR"  
2367..2533  
repeat\_region  
/rpt\_family="(CA)n"  
3148..3270  
repeat\_region  
/rpt\_family="MIR"  
3420..3478  
repeat\_region  
/rpt\_family="MIR"  
complement(3479..3560)  
repeat\_region  
/rpt\_family="AluJc/FLAN"  
3562..3653  
repeat\_region  
/rpt\_family="MIR"  
complement(3772..3831)  
repeat\_region  
/rpt\_family="L2"  
complement(3795..3881)  
repeat\_region  
/rpt\_family="MIR"  
complement(3882..4167)  
repeat\_region  
/rpt\_family="AluJc"  
complement(4168..4291)  
repeat\_region  
/rpt\_family="MIR"  
4853..4933  
repeat\_region  
/rpt\_family="MIR"  
4934..5236  
repeat\_region  
/rpt\_family="AluSx"  
5237..5256  
repeat\_region  
/rpt\_family="MIR"  
5257..5532  
repeat\_region  
/rpt\_family="MULTIH"  
complement(6419..6716)  
repeat\_region  
/rpt\_family="AluJc"  
6752..7164  
repeat\_region  
/rpt\_family="L1MB3"  
7165..7469  
repeat\_region  
/rpt\_family="AluJb"  
7470..7684  
repeat\_region  
/rpt\_family="L1MB3"  
7739..8023  
repeat\_region  
/rpt\_family="AluSg"  
8024..8054  
repeat\_region  
/rpt\_family="(CAA)n"  
8055..8078  
repeat\_region  
/rpt\_family="AluSg"  
8081..8370  
repeat\_region  
/rpt\_family="AluSc"  
8371..8402  
repeat\_region  
/rpt\_family="(CAA)n"  
8663..8725  
repeat\_region  
/rpt\_family="MIR"  
8845..8995  
repeat\_region  
/rpt\_family="CT-rich"  
complement(9002..9274)  
repeat\_region  
/rpt\_family="AluJc"

Query Match 1.7% Score 52; DB 8; Length 149679;  
Best Local Similarity 100.0%; Pred. No. 5.7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAGATTGTCACATGCACTCCAGGCTGGGCAACAGAGCAAGCTGTGTC 3122  
AC012070 AC012070 149776 bp DNA linear PRI 21-ADR-2005  
LOCUS 45739 CAGATTGTCACATGCACTCCAGGCTGGGCAACAGAGCAAGCTGTGTC 45688  
DEFINITION Homo sapiens BAC clone RP11-318K6 from 2, complete sequence.

ACCESSION AC012070 GI:21747558  
 VERSION AC012070.8  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 149776)  
 AUTHORS Teak, A. and Abbott, A.  
 TITLE The sequence of Homo sapiens BAC clone RP11-318K6  
 JOURNAL Unpublished (2001)  
 REFERENCE 2 (bases 1 to 149776)  
 AUTHORS Waterston, R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (19-OCT-1999) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

REFERENCE 3 (bases 1 to 149776)  
 AUTHORS Waterston, R.H.  
 TITLE Direct Submission  
 JOURNAL Submitted (28-FEB-2001) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA  
 REFERENCE 4 (bases 1 to 149776)  
 AUTHORS Waterston, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (09-AUG-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA

REFERENCE 5 (bases 1 to 149776)  
 AUTHORS Waterston, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (07-NOV-2001) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 REFERENCE 6 (bases 1 to 149776)  
 AUTHORS Waterston, R.  
 TITLE Direct Submission  
 JOURNAL Submitted (14-JUL-2002) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
 REFERENCE 7 (bases 1 to 149776)  
 AUTHORS Wilson, R.K.  
 TITLE Direct Submission  
 JOURNAL Submitted (21-APR-2005) Genome Sequencing Center, Washington University School of Medicine, 4444 Forest Park Parkway, St. Louis, MO 63108, USA

On Jul 14, 2002 this sequence version replaced gi:13162544.

----- Genome Center  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: <http://genome.wustl.edu>  
 Contact: [submissions@wustl.edu](mailto:submissions@wustl.edu)  
 ----- Summary Statistics  
 -----  
 Center project name: H\_NH0318K06

# NOTICE:

This sequence was finished as follows unless otherwise noted:  
 all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

MAPPING INFORMATION:  
 Mapping information for this clone was provided by Dr. Wes Warren, Department of Genetics, Washington University, St. Louis MO. For additional information about the map position of this sequence, see <http://genome.wustl.edu>

SOURCE INFORMATION:  
 The RPCI-11 human BAC library was made from the blood of one male

donor, as described by Osogawa, K., Moon, P.Y., Zhao, B., Frengen, E., Tateo, M., Catanesse, J.J. and de Jong, P.J. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.resgen.com>) or Pieter de Jong and coworkers at <http://www.chori.org> VECTOR: pBAC3.6

NEIGHBORING SEQUENCE INFORMATION:  
 The clone sequenced to the left is RP11-335M11, 200 bp overlap; the clone sequenced to the right is RP11-674L1. Actual start of this clone is at base position 150598 of RP11-335M11; actual end is at base position 149776 of RP11-318K6.

There are polymorphic base pair differences in the overlap between the clone RP11-318K6 and RP11-674L1.

## FEATURES

source  
 1..149776  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="2"  
 /clone="RP11-318K6"  
 /clone\_1lb="RPCI-11"  
 9331..9578  
 /note="CpG island (tGC=60.1, o/e=1.05, #CpGs=23)"  
 28997..29314  
 /note="CpG island (tGC=61.0, o/e=0.86, #CpGs=19)"  
 52878..53098  
 /note="CpG island (tGC=60.2, o/e=0.90, #CpGs=18)"  
 ORIGIN

Query Match 1.7%; Score 52; DB 8; Length 149776;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-16;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 2889 GAGCAGGTGATCCTAGCGCCAGAGTTCGAGACCGCTGGCAACAT 2940  
 Db 81593 GAGCAGGTGATCCTAGCGCCAGAGTTCGAGACCGCTGGCAACAT 81644

## RESULT 121

AC018791 150159 bp DNA linear HTG 01-APR-2000  
 LOCUS AC018791  
 DEFINITION Homo sapiens clone RP11-31D10, WORKING DRAFT SEQUENCE, 26 unordered pieces.  
 ACCESSION AC018791  
 VERSION AC018791.3 GI:7382241  
 KEYWORDS HTG; HTGS PHASE1; HTGS\_DRAFT.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 150159)  
 AUTHORS Birren, B., Linton, L., Nusbaum, C. and Lander, B.  
 TITLE Homo sapiens, clone RP11-31D10  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 150159)  
 AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Beckert, R., Bede, F., Boguslavsky, L., Boughalter, B., Brown, A., Burkett, G., Castle, A., Choe, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearlano, K., Dewar, K., Domino, M., Doyle, M., Feneator, J., Ferreira, P., FitzHugh, W., Forrest, C., Gage, D., Galagan, J., Gardina, S., Grant, G., Hages, B., Heaford, A., Horton, L., Howland, J., Johnson, R., Jones, C., Kann, L., Karakas, A., Klein, J., Landers, T., Lechoczky, J., Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K., Meldrum, J., Meneses, L., Morrow, J., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P., Ollivar, T.M., Peterson, K., Plette, N., Plesant, C., Pollara, V., Raymond, C., Riley, R., Rothman, D., Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,

# TITLE JOURNAL COMMENT

Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J., Tirrell,A.,  
Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J., Zimmerman,A.,  
and Zody,M.

## Direct Submission

Submitted (19-DEC-1999) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 1, 2000 this sequence version replaced gi:6592294.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

## Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: <http://www-seq.wi.mit.edu>

Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)

## Project Information

Center project name: L4923

Center clone name: 31 D 10

## Summary Statistics

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 137932 bases at least Q40

Consensus quality: 143842 bases at least Q30

Consensus quality: 146083 bases at least Q20

Insert size: 169000; agarose-efp

Insert size: 147659; sum-of-contigs

Quality coverage: 4.2 in Q20 bases; agarose-efp

Quality coverage: 4.8 in Q20 bases; sum-of-contigs

-----

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 26 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```

1      1239: contig of 1239 bp in length
*      1240      1339: gap of 100 bp
*      1340      3184: contig of 1845 bp in length
*      3185      3284: gap of 100 bp
*      3285      5534: contig of 2250 bp in length
*      5535      5634: gap of 100 bp
*      5635      8315: contig of 2681 bp in length
*      8316      8415: gap of 100 bp
*      8415      10195: contig of 1780 bp in length
*      10196      10295: gap of 100 bp
*      10296      13456: contig of 3161 bp in length
*      13457      13556: gap of 100 bp
*      13557      15457: contig of 1901 bp in length
*      15458      15557: gap of 100 bp
*      15558      17848: contig of 2291 bp in length
*      17849      17948: gap of 100 bp
*      17949      17949: contig of 1945 bp in length
*      17949      19893: gap of 100 bp
*      19894      23457: contig of 3464 bp in length
*      23458      23557: gap of 100 bp
*      23558      26795: contig of 3238 bp in length
*      26796      26895: gap of 100 bp
*      26895      30999: contig of 4104 bp in length
*      31000      31099: gap of 100 bp
*      31100      35140: contig of 4041 bp in length
*      35141      35240: gap of 100 bp
*      35241      39261: contig of 4021 bp in length
*      39262      39361: gap of 100 bp
*      39362      42633: contig of 3272 bp in length
*      42634      42733: gap of 100 bp
*      42734      47924: contig of 5191 bp in length
*      47925      48024: gap of 100 bp
*      48025      52442: contig of 4418 bp in length
*      52443      52542: gap of 100 bp
*      52543      59093: contig of 6551 bp in length

```

## FEATURES

### source

```

*      59094      59193: gap of 100 bp
*      59194      67470: contig of 8277 bp in length
*      67471      67570: gap of 100 bp
*      67571      74426: contig of 6856 bp in length
*      74427      74526: gap of 100 bp
*      74527      87418: contig of 12892 bp in length
*      87419      87518: gap of 100 bp
*      87519      95952: contig of 8434 bp in length
*      95953      96052: gap of 100 bp
*      96053      106060: contig of 10008 bp in length
*      106061      106160: gap of 100 bp
*      106161      121730: contig of 15570 bp in length
*      121731      121830: gap of 100 bp
*      121831      135347: contig of 13517 bp in length
*      135348      135447: gap of 100 bp
*      135448      150159: contig of 14712 bp in length.
*      Location/Qualifiers
1. .150159
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-31D10"
/clone_1b="RP11 Human Male BAC"
1. .1239
/estimated_length=100
1240. .1339
/estimated_length=100
1340. .3184
/estimated_length=100
1385. .3284
/estimated_length=100
3285. .5534
/estimated_length=100
5535. .5634
/estimated_length=100
5635. .8315
/estimated_length=100
8316. .8415
/estimated_length=100
8416. .10195
/estimated_length=100
10196. .10295
/estimated_length=100
10296. .13456
/estimated_length=100
13457. .13556
/estimated_length=100
13557. .15457
/estimated_length=100
15458. .15557
/estimated_length=100
15558. .17848
/estimated_length=100
17849. .17948
/estimated_length=100
17949. .19893
/estimated_length=100
19894. .23457
/estimated_length=100
23458. .23557
/estimated_length=100
23558. .26795
/estimated_length=100
26796. .26895
/estimated_length=100
26896. .30999
/estimated_length=100
31000. .31099
/estimated_length=100
31099. .35140
/estimated_length=100
35141. .39261
/estimated_length=100
39262. .39361
/estimated_length=100
39362. .42633
/estimated_length=100
42634. .42733
/estimated_length=100
42734. .47924
/estimated_length=100
47925. .48024
/estimated_length=100
48025. .52442
/estimated_length=100
52443. .52542
/estimated_length=100
52543. .59093
/estimated_length=100

```

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap

### misc\_feature

gap



```

misc_feature      31100..35140
                    /note="assembly_fragment"
gap               35141..35240
                    /estimated_length=100
misc_feature      35241..39261
                    /note="assembly_fragment"
gap               39262..39361
                    /estimated_length=100
misc_feature      39362..42633
                    /note="assembly_fragment"
                    clone_end=T7
                    vector_size=right"
gap               42634..42733
                    /estimated_length=100
misc_feature      42734..47924
                    /note="assembly_fragment"
gap               47925..48024
                    /estimated_length=100
misc_feature      48025..52442
                    /note="assembly_fragment"
gap               52443..52542
                    /estimated_length=100
misc_feature      52543..59093
                    /note="assembly_fragment"

Query Match      1.7%; Score 52; DB 14; Length 150159;
Best Local Similarity 100.0%; Pred.No.5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3071 CAAGATTGTCACCTGACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122
Db      122542 CAAGATTGTCACCTGACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 122593

RESULT 122
LOCUS      CS086345      150437 bp      DNA      linear      PAT 25-MAY-2005
DEFINITION Sequence 50 from Patent WO2005042786.
ACCESSION  CS086345
VERSION     CS086345.1 GI:66711917
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Homiidae; Homo.
REFERENCE   1
AUTHORS    Harris,C. and Davis,L.
TITLES     Compositions and methods for glioma classification
JOURNAL    Patent: WO 2005042786-A 50 12-MAY-2005;
           Exagen Diagnostics, Inc. (US)
FEATURES
             source
               1..150437
               /organism="Homo sapiens"
               /mol_type="unassigned DNA"
               /db_xref="taxon:9606"

ORIGIN
Query Match      1.7%; Score 52; DB 6; Length 150437;
Best Local Similarity 100.0%; Pred.No.5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3071 CAAGATTGTCACCTGACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122
Db      125751 CAAGATTGTCACCTGACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 125802

RESULT 123
LOCUS      AL138724      150846 bp      DNA      linear      PRI 18-MAY-2005
DEFINITION Human DNA sequence from clone RP11-500C11 on chromosome 6 Contains
             the 5' end of the NUP153 gene for nucleoporin 153kd, the 3' end of
             a gene for a kinesin-like protein (RBKIN) and CpG islands, complete

```

```

sequence.
AL138724
VERSION      AL138724.12 GI:13234804
KEYWORDS     HTG; CpG island; Kinesin; nucleoporin; NUP153; RBKIN.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
             Homiidae; Homo
             1 (bases 1 to 150846)
             Griffiths,C.
REFERENCE    Direct Submission
AUTHORS      Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
TITLES       Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk
JOURNAL      Clone requests: clonerequest@sanger.ac.uk

COMMENT
On Mar 5, 2001 this sequence version replaced gi:13161550.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TrEMBL; Wp: WormPEP; Information
on the WormPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
RP11-500C11 is from the library RPCI-11.2 constructed by the group
of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegas@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.

FEATURES
             source
               1..150846
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"
               /chromosome="6"
               /clone="RP11-500C11"
               /clone_1lb="RPCI-11.2"
               1
               /note="Clone left end; RP11-500C11"
               join(complement(8574..9084),
               complement(AL157776.15:1295..61258),
               complement(AL157776.15:47869..48008),
               complement(AL138824.19:<83780..94906))
               /gene="NUP153"
               /locus_tag="RP11-14H3.6-001"
               join(complement(8974..9084),
               complement(AL157776.15:61036..61258),
               complement(AL157776.15:48162..48410),
               complement(AL157776.15:47869..48008),
               complement(AL157776.15:47545..47673),
               complement(AL157776.15:42070..42186),
               complement(AL157776.15:41933..41977),
               complement(AL157776.15:41615..41668),
               complement(AL157776.15:37879..38025),
               complement(AL157776.15:34658..34710),
               complement(AL157776.15:34293..34419),
               complement(AL157776.15:21803..21940),
               complement(AL157776.15:20447..20545),
               complement(AL157776.15:18707..18784),
               complement(AL157776.15:12579..12704),
               complement(AL157776.15:9793..10410),

```

complement (AL157776.15:5290..5484),  
complement (AL157776.15:1295..2179),  
complement (AL138824.19:94550..94906),  
complement (AL138824.19:93303..93575),  
complement (AL138824.19:85269..85437),  
complement (AL138824.19:83780..84923))  
/gene="NUP153"  
/locus\_tag="RP11-14H3.6-001"  
/product="nucleoporin 153kDa"  
/note="match: ESTs: AA020948 AA280626 AA281956 AA287143  
AA647494 AA987880 AI373182 AI392741 AI568716 AI675749  
AI809345 AI863458 AI918052 AL040791 AL043770 AL048844  
AL597150 AL125074 AL128712 AL137373 AL137888 AL148810  
AL172435 AM269393 AM269401 AM269570 AM502128 AM502296  
AM604841 AM859945 AM859947 AM937856 AM937580 AM972597  
BM529859 BE085871 BE333864 BE381487 BE539035 BE559837  
BE885083 BE886640 BE909229 BE930529 BF120400 BF155241  
BF212928 BF305010 BF702538 BF732659 BF782907 BF843418  
BF844739 BF934823 BF934824 BF934876 BF931889 BF992243  
BG067610 BG080749 BG106248 BG256869 BG259567 BG387515  
BG505279 BG986900 BI040158 BI040160 BI040168 BI052625  
BM090763 F13065 H89627 L44346 R17048 T09409 T75378  
match: cDNAs: AF045567.1 AF434196.1 L06821.1 Z25535.1"  
join(complement(8974..9084),  
complement (AL157776.15:61036..61258),  
complement (AL157776.15:48162..48410),  
complement (AL157776.15:47869..48008),  
complement (AL157776.15:47545..47673),  
complement (AL157776.15:42070..42186),  
complement (AL157776.15:41933..41977),  
complement (AL157776.15:41615..41668),  
complement (AL157776.15:37879..38025),  
complement (AL157776.15:34658..34710),  
complement (AL157776.15:34293..34419),  
complement (AL157776.15:21803..21940),  
complement (AL157776.15:20447..20545),  
complement (AL157776.15:18707..18794),  
complement (AL157776.15:12579..12704),  
complement (AL157776.15:9793..10410),  
complement (AL157776.15:5290..5484),  
complement (AL157776.15:1295..2179),  
complement (AL138824.19:94550..94906),  
complement (AL138824.19:93303..93575),  
complement (AL138824.19:85269..85437),  
complement (AL138824.19:83780..84923))  
/gene="NUP153"  
/locus\_tag="RP11-14H3.6-001"  
/standard\_name="OTTHUMP0000016070"  
/note="match: proteins: O73749 P49790 P49791"  
/codon\_start=1  
/product="nucleoporin 153kDa"  
/protein\_id="CAI40945.1"  
/db\_xref="GI:57208789"  
/db\_xref="GOA:O5TA22"  
/db\_xref="InterPro:IPR001876"  
/db\_xref="UniProt/TrEMBL:O5TA22"  
/translation="MASGAGVGGGGGGKIRRRCHQGEIKRYQGRQGHGILSRVT  
ESVKIVPGMLQRYENKEDVCSCTDTEVPRWPKNDKHLVYADESSNTDTRIT  
PEPAVNTPESTTASNYPLVLTPESLARSHLQVSMLESPALHCOQSTSAFPGS  
SGFSLVKEIKDSTSDHNDISTSGFSRADKIDVLEKSLPLPMGPAERSLS  
SOHTATSSKKPAFNLISAGTSLSPGLGNSILTLTQSDSPFPGKTTVGGAAAVRS  
KLRNTPYQAPVPRQMKAKOLASQSYSTSTRTLIQSEKSSPLADAKRIPSYVSS  
PLNSPLDRSGITIDTFOAGKEKYDSQYFVQMLKPKVSLTNNSVYKRSPLTSGE  
FKTKQRIQNKSTSTGEMKTPGQNNBQKESQSPNFSLPANGLSVGGGGGKMR  
RRTRVAPKPELEBEEVPLPKLSLPTSSSLPFPNFSSEIITVSSPSVQAL  
TKVQMTSSSGSGPMFKFSPSTVKSTEANVLPSSIGFTFVPAKTKELSGSSSL  
BPLSSSAHVTVVNSTCKTPEDCEGPRPAPILKGSVLDILKSGFAPSPIDS  
VAAQPAATSPVYTPRALISSFGSGIGREISKAGSSMCCCTCMLQNTYDNLCTACQ  
AAKSPRPAKQGTGTFPKRSKTTLSAGSTGPKRFXVITGMDCTDLYVNKRBAI  
KCVACETPRGTGCVKRALTLTVVSEALATMTASSSCVTTGTGLGDFKPKISWE  
CAVCCVSNNAEDNKCVCMSKESVSPASSSTVPSVLPSSGSLGLEKFKPRBSMD  
CELCIVQNKADSTKCIACESAKRGTGSGFKGDTSSSSNSAASSFFKGVSSSSSGP

gene  
join (AL023807.6:117981..117993, AL023807.6:142724..142784,  
AL023807.6:160039..160131, AL023807.6:160544..160784,  
AL023807.6:164119..164206, AL023807.6:165704..165838,  
AL023807.6:166672..166784, complement (1140169..140280),  
AL023807.6:139575..139787, complement (1136558..136768),  
complement (133798..133932), complement (1130937..131067),  
complement (1128735..128821), complement (1128465..128631),  
complement (119717..119930), complement (111465..111627),  
complement (108172..108312), complement (1107058..107207),  
complement (102649..102810), complement (1101963..102136),  
complement (99366..99517), complement (97269..97401),  
complement (96946..97092), complement (92569..92607),  
complement (90473..90572), complement (88212..88338),  
complement (86343..86398), complement (83874..83988),  
complement (83427..83603), complement (82289..82381),  
complement (81640..81796), complement (79974..80041),  
complement (77712..77759), complement (76175..76280),  
complement (74605..74756), complement (66392..67643))  
/locus\_tag="RP11-500C11.2-002"  
join (AL023807.6:128627..128752, AL023807.6:29017..29107,  
AL023807.6:117981..117993, AL023807.6:142724..142784,  
AL023807.6:160039..160131, AL023807.6:160544..160784,  
AL023807.6:164119..164206, AL023807.6:165704..165838,  
AL023807.6:166672..166784, complement (1140169..140280),  
AL023807.6:139575..139787, complement (1136558..136768),  
complement (133798..133932), complement (1130937..131067),  
complement (1128735..128821), complement (1128465..128631),  
complement (119717..119930), complement (111465..111627),  
complement (108172..108312), complement (1107058..107207),  
complement (102649..102810), complement (1101963..102136),  
complement (99366..99517), complement (97269..97401),  
complement (96946..97092), complement (92569..92607),  
complement (90473..90572), complement (88212..88338),  
complement (86343..86398), complement (83874..83988),  
complement (83427..83603), complement (82289..82381),  
complement (81640..81796), complement (79974..80041),  
complement (77712..77759), complement (76175..76280),  
complement (74605..74756), complement (66392..67643))  
/locus\_tag="RP11-500C11.2-002"  
join (AL023807.6:128627..128752, AL023807.6:29017..29107,  
AL023807.6:117981..117993, AL023807.6:142724..142784,  
AL023807.6:160039..160131, AL023807.6:160544..160784,  
AL023807.6:164119..164206, AL023807.6:165704..165838,  
AL023807.6:166672..166784, complement (1140169..140280),  
AL023807.6:139575..139787, complement (1136558..136768),  
complement (133798..133932), complement (1130937..131067),  
complement (1128735..128821), complement (1128465..128631),  
complement (119717..119930), complement (111465..111627),  
complement (108172..108312), complement (1107058..107207),  
complement (102649..102810), complement (1101963..102136),  
complement (99366..99517), complement (97269..97401),  
complement (96946..97092), complement (92569..92607),  
complement (90473..90572), complement (88212..88338),  
complement (86343..86398), complement (83874..83988),  
complement (83427..83603), complement (82289..82381),  
complement (81640..81796), complement (79974..80041),  
complement (77712..77759), complement (76175..76280),  
complement (74605..74756), complement (66392..67643))

Query Match 17%; Score 52; DB 8; Length 150846;  
Best Local Similarity 100.0%; Pred. No. 5; 76-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3071 CAAGATTGGCCACGCACTCCAGCGCTGGGCAAGAGAGAAGACTGTGCTC 3122  
DB 4478 CAAGATTGGCCACGCACTCCAGCGCTGGGCAAGAGAGAAGACTGTGCTC 4529

RESULT 124  
AL662800  
LOCUS DEFINITION  
Human DNA sequence from clone Xkbac-24920 on chromosome 6 contains  
the GNL1 gene for guanine nucleotide binding protein-like 1, the  
gene for GATSE protein, the ABCF1 gene for ATP-binding cassette,  
subfamily F, member 1, the PPR10 gene for protein phosphatase 1,  
regulatory subunit 10, the MRPS18 gene for mitochondrial  
ribosomal protein s18b, two novel proteins, a prothymosin alpha  
(PTMA) pseudogene and four CpG islands, complete sequence.  
ACCESSION AL662800  
VERSION AL662800.4 GI:18181758  
KEYWORDS HTG; ABC27; ABC50; ABCF1; ATP-binding cassette, subfamily F, member



```

/db xref="UniProt/TREMBL:Q5STN7"
/translation="MPKRRKONHOPPTQOQPLPERERTGDEDSPIGPSSLGPP
PWANGDGPDKAPAKPNWIPPTCSSEVIMIKLCKHMMGLCHENE"
join(57520..57625,61503..61603)
/genes="PRR3"
/locus tag="XXbac-BPG249D20.5-002"
/standard name="OTTHUMP00000029042"
/notes="match: proteins: P79522 Q91Y5 Q960B9"
/codon_start=1
/product="proline-rich polypeptide 3"
/protein_id="CA118153.1"
/db xref="GI:55961300"
/db xref="UniProt/TREMBL:Q5STN6"
/translation="MPKRRKONHOPPTQOQPLPERERTGDEDSPIAPNWIPTC
PSSSEVIMIKLCKHMMGLCHENE"
join(71568..71735,77584..77630,77993..78088,78248..78377,
78523..78560,78645..78752,80106..80180,80581..80694,
82569..82682,83241..83363,83783..83883,83980..84065,
84367..84506,84593..84741,85435..85518,85733..85857,
86054..86166,86309..86480,86649..86723,86816..86886,
89863..89895,89981..90149,90324..90388,90481..90553,
90710..91487)
/genes="ABCF1"
/locus tag="XXbac-BPG249D20.4-004"
join(71568..71735,77584..77630,77993..78088,78251..78377,
78523..78560,78645..78752,80106..80180,80581..80694,
82569..82682,83241..83363,83783..83883,83980..84065,
84367..84506,84593..84741,85435..85518,85733..85857,
86054..86166,86309..86480,86649..86723,86816..86886,
89863..89895,89981..90149,90324..90388,90481..90553,
90710..91487)
/genes="ABCF1"
/locus tag="XXbac-BPG249D20.4-001"
join(71568..71735,77584..77630,77993..78088,78248..78377,
78523..78560,78645..78752,80106..80180,80581..80694,
82569..82682,83241..83363,83783..83883,83980..84065,
84367..84506,84593..84741,85435..85518,85733..85857,
86054..86166,86309..86480,86649..86723,86816..86886,
89863..89895,89981..90149,90324..90388,90481..90553,
90710..91487)
/genes="ABCF1"
/locus tag="XXbac-BPG249D20.4-001"
join(71568..71735,77584..77630,77993..78088,78248..78377,
78523..78560,78645..78752,80106..80180,80581..80694,
82569..82682,83241..83363,83783..83883,83980..84065,
84367..84506,84593..84741,85435..85518,85733..85857,
86054..86166,86309..86480,86649..86723,86816..86886,
89863..89895,89981..90149,90324..90388,90481..90553,
90710..91487)
/genes="ABCF1"

```

Query Match 1.7%; Score 52; DB 8; Length 151228;  
Best Local Similarity 100.0%; Pred. No. 5.7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3071 CAAGATTGTCACCTGACCTGAGCTGAGCAAGAGCAAGACTGCTGCTC 3122  
12756 CAGATTGTGCACTGCACTGAGCTGAGCAAGAGCAAGACTGCTGCTC 12807

RESULT 125  
AC040905/c 151970 bp DNA 1linear HTG 04-MAY-2000  
LOCUS Homo sapiens chromosome 1 clone RP11-177A12 map 1, WORKING DRAFT  
DEFINITION  
AC040905  
SEQUENCE, 23 unordered pieces.  
AC040905  
VERSION AC040905.2 GI:7705170  
KEYWORDS HTG; HTGS PHAS1; HTGS\_DRAFT.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.  
1 (bases 1 to 151970)  
Birken, B., Linton, L., Nusbaum, C. and Lander, B.  
TITLE Homo sapiens chromosome 1, clone RP11-177A12  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 151970)  
Birken, B., Linton, L., Nusbaum, C., Lander, B., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,

TITLE  
JOURNAL

COMMENT

Boguslavsky, L., Bouhgalter, B., Brown, A., Burkett, G.,  
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,  
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,  
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,  
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
Klein, J., LaRocque, K., Lamazares, R., Lander, C., Lehoczy, J.,  
Levine, R., Lien, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,  
McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McHeeters, R.,  
Meldrum, J., Menus, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,  
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neill, D.C., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,  
Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
Roy, A., Santos, R., Schaner, S., Severy, P., Spencer, B.,  
Stange-Thoman, N., Stojanovic, N., Subramanian, A., Talmas, J.,  
Testaye, S., Theodore, J., Tittell, A., Travers, M., Triggillo, J.,  
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (11-APR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On May 4, 2000 this sequence version replaced gi:7534048.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu

Project Information  
Center project name: 177\_A12  
Center clone name: 177\_A12  
Sequencing vector: M13; M77815; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 139059 bases at least Q40  
Consensus quality: 148027 bases at least Q30  
Insert size: 163000; agarose-fp  
Insert size: 149770; sum-of-contigs  
Quality coverage: 3.7 in Q20 bases; agarose-fp  
Quality coverage: 4.1 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently  
consists of 23 contigs. The true order of the pieces  
is not known and their order in this sequence record is  
arbitrary. Gaps between the contigs are represented as  
runs of N, but the exact sizes of the gaps are unknown.  
This record will be updated with the finished sequence  
as soon as it is available and the accession number will  
be preserved.

1	1024	1123	contig of 1023 bp in length
*	1124	2231	gap of 100 bp
*	2232	2231	contig of 1108 bp in length
*	2232	2231	gap of 100 bp
*	2232	4468	contig of 2137 bp in length
*	4469	4568	gap of 100 bp
*	4569	5885	contig of 1317 bp in length
*	5886	5985	gap of 100 bp
*	5986	8879	contig of 284 bp in length
*	8880	8879	gap of 100 bp
*	8880	10702	contig of 1723 bp in length
*	10703	10802	gap of 100 bp
*	10803	12395	contig of 2193 bp in length
*	12396	13095	gap of 100 bp
*	13096	15614	contig of 2519 bp in length
*	15615	15714	gap of 100 bp
*	15715	19162	contig of 3448 bp in length
*	19163	19262	gap of 100 bp
*	19263	23500	contig of 4238 bp in length

```
* 23501 23500: gap of 100 bp
* 23501 27529: contig of 3923 bp in length
* 27530 27629: gap of 100 bp
* 27630 32861: contig of 5233 bp in length
* 32862 32961: gap of 100 bp
* 32962 37601: contig of 4640 bp in length
* 37602 37701: gap of 100 bp
* 37702 44853: contig of 7153 bp in length
* 44854 44953: gap of 100 bp
* 44954 50843: contig of 5890 bp in length
* 50844 50943: gap of 100 bp
* 50944 60339: contig of 9396 bp in length
* 60340 60439: gap of 100 bp
* 60440 68661: contig of 8222 bp in length
* 68662 68761: gap of 100 bp
* 68762 78468: contig of 9607 bp in length
* 78469 87095: contig of 8627 bp in length
* 87096 87195: gap of 100 bp
* 87196 97272: contig of 10077 bp in length
* 97273 97372: gap of 100 bp
* 97373 112327: contig of 14955 bp in length
* 112328 112427: gap of 100 bp
* 112428 133342: contig of 20915 bp in length
* 133343 133443: gap of 100 bp
* 133443 151970: contig of 18528 bp in length.
```

## FEATURES

## source

```
1. .151970
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/map="1"
/clone="RP11-177A12"
/clone_lib="RPC1-11 Human Male BAC"
1. .1023
/note="assembly_fragment"
1024. .1123
/estimated_length=100
1124. .2231
/note="assembly_fragment"
clone_end:T7
vector_side:right"
2232. .2231
/estimated_length=100
2332. .4468
/note="assembly_fragment"
4469. .4568
/estimated_length=100
4569. .5885
/note="assembly_fragment"
5886. .5985
/estimated_length=100
5986. .8879
/note="assembly_fragment"
8880. .8979
/estimated_length=100
8980. .10702
/note="assembly_fragment"
10703. .10802
/estimated_length=100
10803. .12995
/note="assembly_fragment"
12996. .13095
/estimated_length=100
13096. .15614
/note="assembly_fragment"
15615. .15714
/estimated_length=100
15715. .19162
/note="assembly_fragment"
19163. .19262
/estimated_length=100
```

```
misc_feature 19263. .23500
/note="assembly_fragment"
gap 23501. .23500
/estimated_length=100
misc_feature 23601. .27529
/note="assembly_fragment"
gap 27530. .27629
/estimated_length=100
misc_feature 27630. .32861
/note="assembly_fragment"
gap 32862. .32961
/estimated_length=100
misc_feature 32962. .37601
/note="assembly_fragment"
gap 37602. .37701
/estimated_length=100
misc_feature 37702. .44853
/note="assembly_fragment"
gap 44854. .44953
/estimated_length=100
misc_feature 44954. .50843
/note="assembly_fragment"
gap 50844. .50943
/estimated_length=100
misc_feature 50944. .60339
/note="assembly_fragment"
gap 60340. .60439
/estimated_length=100
misc_feature 60440. .68661
/note="assembly_fragment"
gap 68662. .68761
/estimated_length=100
misc_feature 68762. .78368
/note="assembly_fragment"
gap 78369. .78468
/estimated_length=100
```

Query Match 1.7%; Score 52; DB 14; Length 151970;  
Best Local Similarity 100.0%; Pred.No. 5.7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGGCCACTGCACCTCCAGCCTGGGCAACAGACAGACTGTCTC 3122  
|||||  
Db 100972 CAAGATTGGCCACTGCACCTCCAGCCTGGGCAACAGACAGACTGTCTC 100921

RESULT 126  
AC139547 152040 bp DNA linear HTG 21-FEB-2003  
LOCUS AC139547/c  
DEFINITION Homo sapiens chromosome 15 clone RP13-425K1 map 15, WORKING DRAFT  
SEQUENCE, 24 unordered pieces.  
ACCESSION AC139547  
VERSION AC139547.2 GI:28460953  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.  
REFERENCE 1 (bases 1 to 152040)  
AUTHORS Birren,B., Nuebaum,C. and Lander,B.  
TITLE Homo sapiens chromosome 15, clone RP13-425K1  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 152040)  
AUTHORS Birren,B., Nuebaum,C., Lander,B., Aboueleil,A., Allen,N.,  
Anderson,S., Arachchi,H.M., Barua,N., Bastien,V., Bloom,T.,  
Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J., Choepel,Y.,  
Collamore,A., Cook,A., Cooke,P., Corum,B., DeRellano,K.,  
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,  
Ferreira,P., Fitzgerald,M., Gage,D., Galsgan,J., Gardyna,S.,  
Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,  
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

## TITLE

## JOURNAL

## REFERENCE

## AUTHORS

TITLE  
JOURNAL  
COMMENT

## COMMENT

Lindblad-Toh, K., Liu, G., Lu, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Mhova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Nordu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhkanh, P., Pierre, N., Rachupka, A., Ramsamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Schauer, S., Schupack, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Tsafaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Submitted (05-FEB-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 152040)

Submitted (21-FEB-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Feb 21, 2003 this sequence version replaced gi:28209587.

All repeats were identified using RepeatMasker:

Smith, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

Project Information

Center project name: L29455

Center clone name: 425\_K.1

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 14383 bases at least Q40

Consensus quality: 147021 bases at least Q30

Consensus quality: 148441 bases at least Q20

Insert size: 213000; agarose-fp

Insert size: 149740; sum-of-contigs

Quality coverage: 16.9 in Q20 bases; agarose-fp

Quality coverage: 24.1 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 24 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 210: contig of 210 bp in length

\* 211 310: gap of 100 bp

\* 311 1698: contig of 1388 bp in length

1699 1798: gap of 100 bp

\* 1799 2807: contig of 1009 bp in length

\* 2808 2807: gap of 100 bp

\* 2908 3337: contig of 1030 bp in length

\* 3938 4037: gap of 100 bp

\* 4038 4829: contig of 792 bp in length

\* 4830 4929: gap of 100 bp

\* 4930 5947: contig of 1018 bp in length

\* 5948 6047: gap of 100 bp

\* 6048 7740: contig of 1633 bp in length

\* 7741 7840: gap of 100 bp

\* 7841 9557: contig of 1717 bp in length

\* 9558 9657: gap of 100 bp

\* 9658 10567: contig of 910 bp in length

\* 10568 10567: gap of 100 bp

\* 10668 12556: contig of 1889 bp in length

\* 12557 12656: gap of 100 bp

\* 12657 14292: contig of 1636 bp in length

\* 14293 14392: gap of 100 bp

\* 14393 16890: contig of 2498 bp in length

\* 16891 16890: gap of 100 bp

\* 16891 17263: contig of 273 bp in length

\* 17264 17363: gap of 100 bp

\* 17364 19880: contig of 2517 bp in length

\* 19881 19880: gap of 100 bp

\* 19881 27529: contig of 7549 bp in length

\* 27530 27529: gap of 100 bp

\* 27530 29571: contig of 1942 bp in length

\* 29572 29671: gap of 100 bp

\* 29672 38820: contig of 9149 bp in length

\* 38821 38920: gap of 100 bp

\* 38921 46901: contig of 7961 bp in length

\* 46902 47001: gap of 100 bp

\* 47002 74440: contig of 27439 bp in length

\* 74441 74540: gap of 100 bp

\* 74541 85070: contig of 10530 bp in length

\* 85071 85170: gap of 100 bp

\* 85171 99159: contig of 13989 bp in length

\* 99160 99259: gap of 100 bp

\* 99260 118153: contig of 18894 bp in length

\* 118154 118253: gap of 100 bp

\* 118254 144364: contig of 26111 bp in length

\* 144365 144465: gap of 100 bp

\* 144465 152040: contig of 7576 bp in length.

FEATURES

source

1..152040

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="15"

/map="15"

/clone="RP13-425K1"

/clone\_1lb="RP1-13 Human Female BAC"

1..210

/note="assembly\_fragment"

211..310

/estimated\_length=100

311..1698

/note="assembly\_fragment"

1699..1798

/estimated\_length=100

1799..2807

/note="assembly\_fragment"

2808..2907

/estimated\_length=100

2908..3937

/note="assembly\_fragment"

3938..4037

/estimated\_length=100

4038..4825

/note="assembly\_fragment"

4830..4929

/estimated\_length=100



misc_feature	4930..5947 /note="assembly_fragment"
gap	5948..6047 /estimated_length=100
misc_feature	6048..7740 /note="assembly_fragment"
gap	7741..7840 /estimated_length=100
misc_feature	7841..9557 /note="assembly_fragment"
gap	9558..9657 /estimated_length=100
misc_feature	9658..10567 /note="assembly_fragment"
gap	10568..10667 /estimated_length=100
misc_feature	10668..12556 /note="assembly_fragment"
gap	12557..12656 /estimated_length=100
misc_feature	12657..14292 /note="assembly_fragment"
gap	14293..14392 /estimated_length=100
misc_feature	14393..16890 /note="assembly_fragment"
gap	16891..16990 /estimated_length=100
misc_feature	16991..17263

Query Match 1.7%; Score 52; DB 14; Length 152040;  
Best Local Similarity 100.0%; Pred. No. 5.7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Dy 2889 GAGCGAGTGCATCACTGACGCCGAGATTCCAGACCAGCCTGCACAAT 2940  
|||||  
Db 74741 GAGCGAGTGCATCACCTTGAGGCCAGAGACTTCAGAGCCAGCCTGCCAAT 74690  
|||||

RESULT 127	AC079863	152492 bp	DNA	linear	PRI 21-JUN-2002
LOCUS	AC079863				
DEFINITION	Homo sapiens 3 BAC RP11-9513 (Roswell Park Cancer Institute Human				
ACCESSION	AC079863				
VERSION	AC079863.11	GI:17105254			
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo. 1 (bases 1 to 152492) Muzyly D.M., Adams C., Adio-Oduola B., Ali-osman F.R., Allen C., Albrooks S.L., Amaralunga H.C., Are J.R., Banks T., Barbara J., Berton J., Blmage K., Blankenburg K., Bonnin D., Bouck J., Bowle S., Brieva M., Brown E., Brown M., Bryant N.P., Buhaý C., Butch P., Burkett C., Burrell K.L., Byrd N.C., Carron T.F., Carter M., Cavazos S.R., Chacko J., Chavez D., Chen G., Chen R., Chen Z., Chowdhry I., Christopoulos C., Cleveland C.D., Cox C., Coyle M.D., Dathorne S.R., David R., Davila M.L., Davis C., Davy-carroll L., Dedereich D.A., Delaney K.R., Delgado O., Dem A.L., Ding Y., Dinh H.H., Douhaite K.J., Draper H., Dugan-Rocha S., Durbin K.J., Earnhart C., Edgar D., Edwards C.C., Elina J., Becotto M., Falle T., Ferraguto D., Flagg N., Ford J., Foister P., Frantz P., Gabisi A., Gao J., Garcia A., Garner T., Garza N., Gill R., Gorrell J.H., Guevara W., Gunaratne P., Hale S., Hamilton K., Harris C., Harris C., Hart M., Haylak P., Hawes A., He X., Hernandez J., Hernandez O., Hodgson A., Hogues M., Holloway C., Hollins B., Homes J., Howard S., Huber J., Hulik S., Hume J., Jackson L.E., Jacobson B., Jia Y., Johnson R., Jolivet S., Joudan S., Karlsson E., Kelly S., Khan U., King L., Korvah J., Kovari C., Kratovic J., Kureshi A., Landry N., Leal B., Lewis L.C.,				
REFERENCE	AUTHORS				

Lewis, L., Li, J., Li, Z., Lichargos, O., Lieu, C., Liu, J., Liu, W.,  
 Lougheed, R., Lozano, R. J., Lu, X., Lucier, A., Lucier, R., Luna, R.,  
 Ma, J., Maheshwari, M., Mapa, P., Martin, R., Martindale, A.,  
 Martinez, E., Massey, E., Mathew, E., McLeod, M. P., Meador, M.,  
 Mel, G., Metzger, M., Miner, G., Miner, Z., Mitchell, T., Mohabbat, K.,  
 Moore, S., Morgan, M., Moorish, T., Morris, S., Moser, M., Neal, D.,  
 Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N.,  
 Nickerson, E., Nwokkenkwo, S., Ogih, M., Okunolu, G., Ogunye, N.,  
 Oyedokun, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L.,  
 Pickens, R., Primus, E., Pu, L., Quiles, M., Ren, Y., Rivas, M.,  
 Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S.,  
 Scott, G., Shen, H., Shoshchik, N., Sisson, I., Sodergren, E.,  
 Sornalik, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A.,  
 Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H., Tansey, J., Taylor, C.,  
 Taylor, T., Telford, B., Thomas, N., Thomas, S., Umanak, K., Vasquez, L.,  
 Vera, Y., Villalon, D., Vinson, R., Wall, R., Wang, S., Ward-Moore, S.,  
 Warren, R., Washington, C., Watlington, S., Williams, G.,  
 Williamson, A., Wleczek, R., Wooden, S., Worley, K., Wu, C., Wu, Y.,  
 Wu, Y. F., Zhou, J., Zorrilla, S., Naylor, S. L., Weinstein, G., and  
 Gibbs, R.  
 Direct Submission  
 2 (bases 1 to 152492)  
 Worley, K. C.  
 Direct Submission  
 Submitted (15-SEP-2000) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 3 (bases 1 to 152492)  
 Worley, K. C.  
 Direct Submission  
 Submitted (05-DEC-2001) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 4 (bases 1 to 152492)  
 Worley, K. C.  
 Direct Submission  
 Submitted (07-MAR-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 5 (bases 1 to 152492)  
 Worley, K. C.  
 Direct Submission  
 Submitted (28-MAR-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 6 (bases 1 to 152492)  
 Worley, K. C.  
 Direct Submission  
 Submitted (21-JUN-2002) Human Genome Sequencing Center, Department  
 of Molecular and Human Genetics, Baylor College of Medicine, One  
 Baylor Plaza, Houston, TX 77030, USA  
 On Nov 27, 2001 this sequence version replaced gi:16874814.  
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
[gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

CLONE LENGTH: This sequence does not necessarily represent the  
 entire insert of this clone. Overlapping regions of clones are only  
 sequenced and submitted once, so the sequence for the remainder of  
 the insert may be found in the record for the adjacent clones.  
 Overlapping clones are noted at the beginning and end of the  
 Features listing.

ANNOTATION OF FEATURES:  
 STS are identified using ePCR (Genome Res. 7:541-550) searches  
 of a local database that includes entries from dbSTS, GDB, and  
 local mapping efforts.  
 Repeats are identified using RepeatMasker (A. Smit and P. Green  
 unpublished.) for Human and Mouse sequences.  
 Genes and Region of sequence similarity are identified by BLAST  
 (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the  
 EST and cDNA sequences. Genes demonstrate at least two exons  
 flanked by consensus splice sites that maintained sequence

continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL: <http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

## QUALSTAT-REPORT.

FEATURES  
source location/Qualifiers

1. 152492  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/clone="RP11-95L3"  
1. 2000  
/note="overlaps bases 166235..168234 of clone AC092952"  
/function="clone overlap"

repeat\_region /rpt\_family="MIR"  
complement(2551..2690)  
repeat\_region /rpt\_family="L1PA13"  
2820..3079  
repeat\_region /rpt\_family="L2"  
3321..3367  
repeat\_region /rpt\_family="MADE1"  
4166..4312  
repeat\_region /rpt\_family="GA-rich"  
complement(4369..4617)  
repeat\_region /rpt\_family="MIR"  
complement(4924..5221)  
repeat\_region /rpt\_family="AluUc"  
complement(5853..6280)  
repeat\_region /rpt\_family="L1"  
6281..6339  
repeat\_region /rpt\_family="Tigge2"  
6346..6785  
repeat\_region /rpt\_family="Tigge2"  
complement(6813..6882)  
repeat\_region /rpt\_family="PAM"  
6924..7222  
repeat\_region /rpt\_family="LIM3e"  
complement(7239..7419)  
repeat\_region /rpt\_family="7S18RNA"  
7420..7622  
repeat\_region /rpt\_family="Tigge1"  
complement(7627..7975)  
repeat\_region /rpt\_family="LIM2"  
complement(7990..8633)  
repeat\_region /rpt\_family="LIM4"  
8637..8665  
repeat\_region /rpt\_family="AT-rich"  
8672..8692  
repeat\_region /rpt\_family="AT-rich"  
8730..8820  
repeat\_region /rpt\_family="MLT1B2"  
8833..8861  
repeat\_region /rpt\_family="AT-rich"  
complement(8862..8974)  
repeat\_region /rpt\_family="AluUcX"  
9078..9169  
repeat\_region /rpt\_family="MLT1B2"  
9272..9486

repeat\_region /rpt\_family="MLT1B2"  
complement(9489..10203)  
repeat\_region /rpt\_family="LIM4"  
10423..10584  
repeat\_region /rpt\_family="LIMC"  
10891..11733  
repeat\_region /rpt\_family="LIMC"  
11835..12106  
repeat\_region /rpt\_family="AluUc"  
12207..12566  
repeat\_region /rpt\_family="LIMC"  
complement(12623..12990)  
repeat\_region /rpt\_family="MLT1D"  
complement(13033..14014)  
repeat\_region /rpt\_family="LIM4"  
complement(14015..14326)  
repeat\_region /rpt\_family="AluUc"  
complement(14327..14539)  
repeat\_region /rpt\_family="LIM4"  
14540..14823  
repeat\_region /rpt\_family="Tigge3b"  
complement(14824..15113)

Query Match 1.7%; Score 52; DB 8; Length 152492;  
Best Local Similarity 100.0%; Pred. No. 5.7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTCACCTGCACCTCCAGCCTGGGCAACAGACAGACTGTCTCTC 3122  
DB 151355 CAAGATTGTCACCTGCACCTCCAGCCTGGGCAACAGACAGACTGTCTCTC 151406

RESULT 128  
AC064817/c 157135 bp DNA HTG 16-JUL-2000  
LOCUS  
DEFINITION Homo sapiens chromosome 17 clone RP11-46122, WORKING DRAFT  
SEQUENCE, 27 unordered pieces.

AC064817  
AC064817.4 GI:8980939  
VERSION  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 157135)  
Waterston, R.H.

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
2 (bases 1 to 157135)  
Waterston, R.H.  
Direct Submission  
Submitted (22-Apr-2000) Genome Sequencing Center, Washington  
University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
MO 63108, USA

COMMENT  
On Jul 9, 2000 this sequence version replaced gi:7770662.

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUSGC  
Web site: <http://genome.wustl.edu/gsc/index.shtml>  
----- Project Information -----  
Center project name: H\_NH0046122  
----- Summary Statistics -----  
Sequencing vector: M13; 100%  
Sequencing vector: Plasmid; 0%  
Chemistry: Dye-terminator Big Dye; 0% of reads  
Chemistry: Dye-terminator Big Dye; 0% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 141363 bases at least Q40  
Consensus quality: 146787 bases at least Q30  
Consensus quality: 149152 bases at least Q20  
Insert size: 165000; agarose-ftp

Insert size: 154535; sum-of-contigs  
Quality coverage: 3.28 in Q20 bases; agarose-fp  
Quality coverage: 3.81 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 27 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1120: contig of 1120 bp in length  
\* 1121 1220: gap of unknown length  
\* 1221 2618: contig of 1398 bp in length  
\* 2619 2718: gap of unknown length  
\* 2719 4808: contig of 2090 bp in length  
\* 4809 4908: gap of unknown length  
\* 4909 6698: contig of 1790 bp in length  
\* 6699 6798: gap of unknown length  
\* 6799 9212: contig of 2414 bp in length  
\* 9213 9312: gap of unknown length  
\* 9313 11332: contig of 2020 bp in length  
\* 11333 11432: gap of unknown length  
\* 11433 14518: contig of 3086 bp in length  
\* 14519 14618: gap of unknown length  
\* 14619 18220: contig of 3602 bp in length  
\* 18221 18320: gap of unknown length  
\* 18321 21383: contig of 3063 bp in length  
\* 21384 21483: gap of unknown length  
\* 21484 25924: contig of 4441 bp in length  
\* 25925 26024: gap of unknown length  
\* 26025 29223: contig of 3199 bp in length  
\* 29224 29323: gap of unknown length  
\* 29324 33796: contig of 4473 bp in length  
\* 33797 33896: gap of unknown length  
\* 33897 37526: contig of 3630 bp in length  
\* 37527 37626: gap of unknown length  
\* 37627 42125: contig of 4499 bp in length  
\* 42126 42225: gap of unknown length  
\* 42226 47482: contig of 5257 bp in length  
\* 47483 47582: gap of unknown length  
\* 47583 53156: contig of 5574 bp in length  
\* 53157 53256: gap of unknown length  
\* 53257 59938: contig of 6682 bp in length  
\* 59939 60038: gap of unknown length  
\* 60039 66057: contig of 6019 bp in length  
\* 66058 66157: gap of unknown length  
\* 66158 74668: contig of 8511 bp in length  
\* 74669 74768: gap of unknown length  
\* 74769 82448: contig of 7680 bp in length  
\* 82449 82548: gap of unknown length  
\* 82549 91047: contig of 8499 bp in length  
\* 91048 91147: gap of unknown length  
\* 91148 99649: contig of 8502 bp in length  
\* 99650 99749: gap of unknown length  
\* 99750 109752: contig of 10003 bp in length  
\* 109753 109852: gap of unknown length  
\* 109853 120015: contig of 10163 bp in length  
\* 120016 120116: gap of unknown length  
\* 120117 131890: contig of 11775 bp in length  
\* 131891 131990: gap of unknown length  
\* 131991 141546: contig of 9556 bp in length  
\* 141547 157135: contig of 15489 bp in length.  
\* 141647

## FEATURES

## source

1.157135  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/clone="RP11-46122"  
1.1120

## misc\_feature

1.1120

gap  
misc\_feature

/note="assembly\_name:Contig15"  
1121..1220  
/estimated\_length=unknown  
1221..2618  
/note="assembly\_name:Contig16  
clone end:SP6  
vector side:left"

gap  
misc\_feature

/estimated\_length=unknown  
2619..2718  
/estimated\_length=unknown  
2719..4808  
/note="assembly\_name:Contig17"

gap

/estimated\_length=unknown  
4809..4908  
/estimated\_length=unknown  
4909..6698

gap  
misc\_feature

/note="assembly\_name:Contig18"  
6699..6798  
/estimated\_length=unknown  
6799..9212  
/note="assembly\_name:Contig19"

gap  
misc\_feature

/estimated\_length=unknown  
9213..9312  
/estimated\_length=unknown  
9313..11332  
/note="assembly\_name:Contig20"

gap  
misc\_feature

/estimated\_length=unknown  
11333..11432  
/estimated\_length=unknown  
11433..14518  
/note="assembly\_name:Contig21"

gap  
misc\_feature

/estimated\_length=unknown  
14519..14618  
/estimated\_length=unknown  
14619..18220  
/note="assembly\_name:Contig22"

gap  
misc\_feature

/estimated\_length=unknown  
18221..18320  
/estimated\_length=unknown  
18321..21383  
/note="assembly\_name:Contig23"

gap  
misc\_feature

/estimated\_length=unknown  
21384..21483  
/estimated\_length=unknown  
21484..25924  
/note="assembly\_name:Contig24"

gap  
misc\_feature

/estimated\_length=unknown  
25925..26024  
/estimated\_length=unknown  
26025..29223  
/note="assembly\_name:Contig25"

gap  
misc\_feature

/estimated\_length=unknown  
29224..29323  
/estimated\_length=unknown  
29324..33796  
/note="assembly\_name:Contig26"

gap  
misc\_feature

/estimated\_length=unknown  
33797..33896  
/estimated\_length=unknown  
33897..37526  
/note="assembly\_name:Contig27"

gap  
misc\_feature

/estimated\_length=unknown  
37527..37626  
/estimated\_length=unknown  
37627..42125  
/note="assembly\_name:Contig28"

gap  
misc\_feature

/estimated\_length=unknown  
42126..42225  
/estimated\_length=unknown  
42226..47482  
/note="assembly\_name:Contig29"

gap  
misc\_feature

/estimated\_length=unknown  
47483..47582  
/estimated\_length=unknown  
47583..53156  
/note="assembly\_name:Contig30"

gap  
misc\_feature

/estimated\_length=unknown  
53157..53256  
/estimated\_length=unknown  
53257..59938  
/note="assembly\_name:Contig31"

gap  
misc\_feature

/estimated\_length=unknown  
59939..60038  
/estimated\_length=unknown  
60039..66057  
/note="assembly\_name:Contig32"

gap

/estimated\_length=unknown  
66058..66157  
/estimated\_length=unknown

```

misc_feature      66158..74668
                    /note="assembly_name:Contig33"
gap               74669..74768
                    /estimated_length=unknown
misc_feature      74769..82448
                    /note="assembly_name:Contig34"
gap               82449..82548
                    /estimated_length=unknown
misc_feature      82549..91047
                    /note="assembly_name:Contig35"
gap               91048..91147
                    /estimated_length=unknown
misc_feature      91148..99649
                    /note="assembly_name:Contig36"
gap               99650..99749

Query Match
Best Local Similarity 1.7%; Score 52; DB 14; Length 157135;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3071 CAAGATTGCGCATCTGCACCTGCGGCAACAGACGACGCTGTCTC 3122
Db      147583 CAAGATTGCGCATCTGCACCTGCGGCAACAGACGACGCTGTCTC 147532

RESULT 129
LOCUS      AC008677      160929 bp      DNA      linear      PRI 01-JAN-2002
DEFINITION Homo sapiens chromosome 5 clone CTB-47B8, complete sequence.
ACCESSION  AC008677
VERSION     AC008677.5 GI:18030123
KEYWORDS    HTG.
SOURCE      Homo sapiens (human)
  ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo.
REFERENCE   1 (bases 1 to 160929)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE     Direct Submission
  JOURNAL   Unpublished
  COMMENT   2 (bases 1 to 160929)
            Direct Submission
            DOE Joint Genome Institute.
REFERENCE   2 (bases 1 to 160929)
  JOURNAL   Direct Submission
  TITLE     Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
            3 (bases 1 to 160929)
            DOE Joint Genome Institute and Stanford Human Genome Center.
            Direct Submission
            Submitted (01-JAN-2002) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
            On Jan 1, 2002 this sequence version replaced gi:7709258.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing completed at Stanford Human Genome Center
            www.shgc.stanford.edu
            Quality: Phrap Quality >=40 99.9% of Sequence;
            Estimated Total Number of Errors is 0.1.
FEATURES
  source
    1..160929
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="5"
    /clone="CTB-47B8"

ORIGIN
Query Match      1.7%; Score 52; DB 8; Length 160929;
Best Local Similarity 100.0%; Pred. No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2889 GAGGACGTGATCATCTGAGGCGCAGAGTTGAGACGACCTGCGCAACAT 2940
|||||

```

```

Db      92839 GAGGACGTGATCATCTGAGGCGCAGAGTTGAGACGACCTGCGCAACAT 92890

RESULT 130
AC027706/C
LOCUS      AC027706      161970 bp      DNA      linear      HTG 02-MAY-2000
DEFINITION Homo sapiens chromosome 12 clone RP11-665C13 map 12, WORKING DRAFT
            SEQUENCE, 49 unordered pieces.
ACCESSION  AC027706
VERSION     AC027706.2 GI:7677821
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens (human)
  ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo.
REFERENCE   1 (bases 1 to 161970)
  AUTHORS   Birren, B., Linton, L., Nusbaum, C. and Lander, B.
  TITLE     Homo sapiens chromosome 12, clone RP11-665C13
  JOURNAL   Unpublished
  COMMENT   2 (bases 1 to 161970)
            Birren, B., Linton, L., Nusbaum, C., Lander, B., Abraham, H., Allen, N.,
            Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
            Boguslavsky, L., Borkhagalter, B., Brown, A., Burkett, G.,
            Campopiano, A., Castle, A., Choquel, Y., Colangelo, M., Collins, S.,
            Collamore, A., Cooke, P., DeRellano, K., Dewar, K., Diaz, J.S.,
            Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,
            Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
            Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,
            Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
            Klein, J., Labrecque, K., Lamazares, R., Lander, T., Lehotzky, J.,
            Levine, R., Lien, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,
            McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McPheters, R.,
            Melatim, J., Meneses, L., Milhova, T., Miranda, C., Mlenga, V., Morrow, J.,
            Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
            O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,
            Pisanu, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
            Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
            Stange-Thomann, N., Stojanovic, N., Subramanian, A., Taimas, J.,
            Testaye, S., Theodore, J., Tittell, A., Travers, M., Triggillo, J.,
            Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
            Young, G., Zainoun, J., Zimmer, A. and Zody, M.
            Direct Submission
            Submitted (01-Apr-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
            On May 2, 2000 this sequence version replaced gi:7582510.
            All repeats were identified using RepeatMasker:
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WISR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information
Center project name: 16713
Center clone name: 665_C13
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.96071
Consensus quality: 140943 bases at least Q40
Consensus quality: 150450 bases at least Q40
Consensus quality: 154599 bases at least Q20
Insert size: 165000; agarose-fp
Insert size: 157170; sum-of-coverage
Quality coverage: 3.1 in Q20 bases; agarose-fp
Quality coverage: 3.3 in Q20 bases; sum-of-coverage

-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 49 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as

```

*	runs of N, but the exact sizes of the gaps are unknown.
*	This record will be updated with the finished sequence
*	as soon as it is available and the accession number will
*	be preserved.
*	1
*	1052: contig of 1052 bp in length
*	1053
*	1152: gap of 100 bp
*	1153
*	2378: contig of 1226 bp in length
*	2379
*	2479: gap of 100 bp
*	3812: contig of 1334 bp in length
*	3813
*	3912: gap of 100 bp
*	3913
*	5368: contig of 1456 bp in length
*	5369
*	5468: gap of 100 bp
*	5469
*	6519: contig of 1051 bp in length
*	6520
*	6619: gap of 100 bp
*	6620
*	7552: contig of 1233 bp in length
*	7553
*	7952: gap of 100 bp
*	7953
*	8464: contig of 512 bp in length
*	8465
*	8564: gap of 100 bp
*	8565
*	10347: contig of 1783 bp in length
*	10348
*	10447: gap of 100 bp
*	12145: contig of 1698 bp in length
*	12146
*	12245: gap of 100 bp
*	14006: contig of 1761 bp in length
*	14007
*	14106: gap of 100 bp
*	14306: contig of 1252 bp in length
*	15358
*	15458: gap of 100 bp
*	15459
*	17152: contig of 1694 bp in length
*	17153
*	17252: gap of 100 bp
*	17253
*	18800: contig of 1548 bp in length
*	18801
*	18900: gap of 100 bp
*	18901
*	20231: contig of 1391 bp in length
*	20292
*	20311: gap of 100 bp
*	20392
*	21889: contig of 1498 bp in length
*	21890
*	21969: gap of 100 bp
*	21990
*	24355: contig of 2366 bp in length
*	24356
*	24455: gap of 100 bp
*	24456
*	26757: contig of 2302 bp in length
*	26758
*	26857: gap of 100 bp
*	26858
*	28317: contig of 1460 bp in length
*	28318
*	28417: gap of 100 bp
*	30761: contig of 2364 bp in length
*	30782
*	30881: gap of 100 bp
*	30882
*	32789: contig of 1908 bp in length
*	32790
*	32889: gap of 100 bp
*	32890
*	35214: contig of 2335 bp in length
*	35215
*	35314: gap of 100 bp
*	35315
*	37181: contig of 1867 bp in length
*	37182
*	37281: gap of 100 bp
*	37282
*	39642: contig of 2361 bp in length
*	39643
*	39782: gap of 100 bp
*	39783
*	41978: contig of 2226 bp in length
*	41979
*	42078: gap of 100 bp
*	42079
*	44963: contig of 2885 bp in length
*	44964
*	45063: gap of 100 bp
*	45064
*	47255: contig of 2202 bp in length
*	47256
*	47365: gap of 100 bp
*	47366
*	49887: contig of 2482 bp in length
*	49888
*	49917: gap of 100 bp
*	49918
*	52474: contig of 2527 bp in length
*	52475
*	52574: gap of 100 bp
*	52575
*	54767: contig of 2193 bp in length
*	54768
*	54867: gap of 100 bp
*	54868
*	57998: contig of 3071 bp in length
*	57939
*	58038: gap of 100 bp
*	58039
*	60538: contig of 2500 bp in length
*	60539
*	60638: gap of 100 bp
*	60639
*	63422: contig of 2784 bp in length
*	63423
*	63522: gap of 100 bp
*	63523
*	66038: contig of 2516 bp in length
*	66039
*	66138: gap of 100 bp
*	66139
*	69185: contig of 3047 bp in length
*	69186
*	69255: gap of 100 bp
*	69256
*	72311: contig of 3036 bp in length

[illegible]

a major histocompatibility complex class I-related polypeptide sequence B (MICB) pseudogene, a thymopoietin (TMPO) pseudogene, a succinate-CoA ligase, ADP-forming, beta subunit (SUCLA2) pseudogene, a RAN, member RAS oncogene family pseudogene, the HLA-E gene for major histocompatibility complex, class I, E and six CpG islands, complete sequence.

AL662822  
AL662822.4 GI:18650691  
HTG: A-BETA: ANA24: CpG island; HLA-6.2; HLA-E; major histocompatibility complex class I-related p; major histocompatibility complex, class I, E; MICB; RAN; succinate-CoA ligase; SUCLA2; TC4; thymopoietin; TMPO; TP.

Homo sapiens (human)

SOURCE  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 162948)

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Johnson, C.  
Direct Submissions  
Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk  
Clone requests: clonerequest@sanger.ac.uk  
On Feb 11, 2002 this sequence version replaced gi:18135278.  
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:  
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at  
http://www.sanger.ac.uk/Projects/C\_elegans/wormpep  
Xxbac-17K10 is from a CHORI-502 human bac - COX cell line library VECTOR: PTARBAC2.1

This sequence was generated from part of bacterial clone configs constructed by the MHC Haplotype Consortium and collaborators.  
Further information can be found at  
http://www.sanger.ac.uk/HGP/Ch6/MHC

----- Genome Center  
Center: Wellcome Trust Sanger Institute  
Center code: SC  
Web site: http://www.sanger.ac.uk  
Contact: vegas@sanger.ac.uk

-----  
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

Location/Qualifiers  
1. 162948  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="6"  
/clone="Xxbac-17K10"  
/clone\_1fb="CHORI-502"  
1  
/note="Clone\_left\_end: Xxbac-17K10"  
28257  
/note="Clone\_right\_end: Xxbac-22J1"  
join(32973..33223,33490..33775,34377..34894,37463..37598)  
/locus\_tag="Xxbac-BCX17K10.1-001"  
/locus\_tag="Xxbac-BCX17K10.1-001"  
/pseudo  
join(32973..33223,33490..33775,34377..34894,37463..37598)  
/locus\_tag="Xxbac-BCX17K10.1-001"  
/note="match: proteins: O98234 O98266 O98269 Q8WLO0 Q8WLO1 Q8WLO6 Q95HA2 Q9BCR4 Q9BCR5 Q9GIZ5 Q9TPU8"  
/pseudo  
/codon\_start=1  
/product="MHC class I polypeptide-related sequence C"  
79963  
/note="Clone\_left\_end: Xxbac-98J21"  
join(84827..85246,85498..86253)  
gene

/locus\_tag="Xxbac-BCX17K10.3-001"  
/pseudo  
join(84827..85246,85498..86253)  
/locus\_tag="Xxbac-BCX17K10.3-001"  
/note="match: proteins: AAC2390 O73766 P42167 Q61028 Q61030 Q61031 Q61032"  
/pseudo  
/codon\_start=1  
/product="thymopoietin-like 1"  
87207..88544  
/locus\_tag="Xxbac-BCX17K10.5-001"  
/pseudo  
87207..88544  
/locus\_tag="Xxbac-BCX17K10.5-001"  
/note="match: proteins: O97580 P53312 P53568 Q8W998 Q9P2R7 Q9Y470 Q9Y137 Q9Z219"  
/pseudo  
/codon\_start=1  
/product="succinate-CoA ligase, ADP-forming, beta subunit pseudogene"  
104251..104901  
/locus\_tag="Xxbac-BCX17K10.2-001"  
/pseudo  
104251..104901  
/locus\_tag="Xxbac-BCX17K10.2-001"  
/note="match: proteins: AAH14518 AAH14829 AAM33416 O76173 B17080 P32835 P35292 P42558 P91857 Q22908 Q61820 Q91BE8 Q9XU45 Q9YGC0"  
/pseudo  
/codon\_start=1  
/product="RAN, member RAS oncogene family pseudogene 1"  
107752..112507  
/gene="HLA-B"  
/locus\_tag="Xxbac-BCX17K10.4-001"  
join(107752..107905,108036..108305,108550..108825,109447..109722,109847..109963,110714..110746,110851..110893,111059..112507)  
/gene="HLA-B"  
/locus\_tag="Xxbac-BCX17K10.4-001"  
/product="major histocompatibility complex, class I, E"  
/note="match: ESTs: BG028524.1 BG756822.1 BM791769.1 BM917068.1  
match: CDNA: AJ293263.1 AJ293264.1 BC002578.1 M20022.1  
join(107842..107905,108036..108305,108550..108825,109447..109722,109847..109963,110714..110746,110851..110893)  
/gene="HLA-B"  
/locus\_tag="Xxbac-BCX17K10.4-001"  
/standard\_name="OTTHUMP0000029626"  
/note="match: proteins: Q9G1Y7 Q9G1Y8"  
/codon\_start=1  
/product="major histocompatibility complex, class I, E"  
/protein\_id="CA117548.1"  
/db\_xref="GI:55961319"  
/db\_xref="InterPro:IPR001039"  
/db\_xref="InterPro:IPR003006"  
/db\_xref="InterPro:IPR003597"  
/db\_xref="InterPro:IPR007110"  
/translation="VVDGTLILSLBALITGTWAGSHSLKTFHTSVSRPGSGPRFI SGYVDVDDQFVTFDNDASPRVPRPAPWMEQSGSYWDRSTASADTAQIFVNLATL RGVYNQSESGFTVLOMHGCELGPRPRFGRGEQSYAYDGKDYLTINEDLRSTAVADTA AQISBOKNLSASBAHORAYLEDTCEVMHKLKLBGKETLLHLPKTHVTHPISDH BAIIRCKMLAGPYPARITLTWODGSGHODPELVETPRAGDGTPOKMAVVVPSGEO RYTCVQHGEGPEPTLTMKRPAASQPTITIVGIIAGLVLGSVVGAVVAIVTRKKS GKGGSYSIAEWSDAQSGESHS1L"  
162948  
/note="Clone\_right\_end: Xxbac-17K10"

## ORIGIN

Query Match 1.7%; Score 52; DB 8; Length 162948;  
Best Local Similarity 100.0%; Freq. No. 5.7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



Db	3071	CAAGATTGTGCACATGCACCTCAGGCTGGGCAACAGACGACATCTCTCTC	3122
RESULT 132	AC108670	163157 bp	DNA linear PRI 28-DEC-2002
LOCUS	AC108670/c		
DEFINITION	Homo sapiens 3 BAC RP11-218A22 (Roswell Park Cancer Institute Human BAC library) complete sequence.		
ACCESSION	AC108670		
VERSION	AC108670.10	GI:23555671	
KEYWORDS	HTG.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	1 (bases 1 to 163157)		
	Murphy,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C., Albrooks,S.L., Amaralunge,H.C., Are,J.R., Ayele,M., Banks,T., Barberia,J., Benton,J., Blmage,K., Blankenburg,K., Bonani,D., Bouck,J., Boyle,S., Briteva,M., Brown,E., Brown,M., Bryant,N.P., Bunay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C., Carron,T.R., Carter,M., Cazazos,S.R., Chacko,U., Chavez,D., Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C., Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R., Dayla,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A., Delaney,K.R., Delgado,O., Dem,A.L., Ding,Y., Dint,H.H., Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J., Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M., Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P., Gabisi,A., Gao,U., Garcia,A., Garner,T., Garza,N., Gill,R., Gorrell,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K., Harris,C., Harris,K., Hart,M., Havlik,P., Hawes,A., He,X., Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Hollway,C., Hollins,B., Homsi,F., Howard,S., Huber,J., Hulyk,S., Hune,J., Jackson,L.E., Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudan,S., Karlsson,E., Kelly,S., Khan,U., Kling,L., Koryak,J., Kovari,C., Kretovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L., Li,J., Li,Z., Lichtarge,O., Ilen,C., Liu,J., Liu,W., Lonsheed,R., Lozano,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E., Massey,E., Mahoney,E., Mcleod,M.P., Medor,M., Mel,G., Metzner,M., Miner,G., Miner,Z., Mitchell,T., Mohabot,K., Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D., Nelson,D., Newton,J., Newton,S., Nguyen,A., Nguyen,N., Nickerson,G., Nwokkenwo,S., Ogbu,M., Okunolu,G., Oragunye,N., Oyedele,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L., Piekens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Ruiz,S., Savery,G., Scherer,S., Scott,G., Shen,H., Shoshetari,N., Slason,I., Sodegryn,E., Sonalik,T., Sparks,A., Stalley,H., Stone,H., Sutton,A., Stavek,A., Tabori,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmari,K., Vazquez,L., Vera,V., Villalon,D., Vnson,R., Wang,Q., Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S., Williams,G., Williamson,A., Wiczek,R., Wooden,S., Worley,K., Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorilla,S., Naylor,S.L., Weinstein,G. and Gibbs,R.		
TITLE	Direct Submission		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 163157)		
AUTHORS	Worley,K.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (31-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA		
REFERENCE	3 (bases 1 to 163157)		
AUTHORS	Worley,K.C.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One		

```

REFERENCE
AUTHORS
TITLE
JOURNAL
4 (bases 1 to 163157)
Baylor Plaza, Houston, TX 77030, USA
Morley,K.C.
Direct Submission
Submitted (30-SEP-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
5 (bases 1 to 163157)
Morley,K.C.
Direct Submission
Submitted (01-OCT-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
6 (bases 1 to 163157)
Morley,K.C.
Direct Submission
Submitted (28-DEC-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Sep 30, 2002 this sequence version replaced gi:20514653.
INFORMATION: http://www.hgsc.bcm.tmc.edu/ or email
gc-help@bcm.tmc.edu

COMMENT
CLONE LENGTH: This sequence does not necessarily represent the
entire insert of this clone. Overlapping regions of clones are only
sequenced and submitted once, so the sequence for the remainder of
the insert may be found in the record for the adjacent clones.
Overlapping clones are noted at the beginning and end of the
Features listing.

ANNOTATION OF FEATURES:
STS are identified using ePCR (Genome Res. 7:541-550) searches
of a local database that includes entries from dbSTS, GDB, and
local mapping efforts.
Repeats are identified using RepeatMasker (A. Smit and P. Green,
unpublished.) for Human and Mouse sequences.
Genes and Region of sequence similarity are identified by BLAST
(Nuc. Acids Res. 25:3389-3402) similarly (expect < 1e-34) to the
EST and cDNA sequences. Genes demonstrate at least two exons
flanked by consensus splice sites that maintained sequence
continuity across the splice junctions. Sequences that are not
identical matches are annotated as similar.

SEQUENCING READ COVERAGE: sequencing is completed to a minimum
standard of double strand coverage with a minimum of 2 clones and 2
reads with no ambiguities or 2 chemistries with a minimum of 2
clones and 3 reads with no ambiguities. If the sequence quality for
a region does not meet this standard, it will be indicated in the
annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
standards - estimated error rate less than 1 per 10,000 bases.
Reports of lowest quality individual bases and measures of base
quality are listed below. Description of the metrics can be found
at URL:
http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht
ml.

FEATURES
source
location/Qualifiers
1..163157
/organelle="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="X"
/clone="RP11-218A22"
1..234
/rpc_family="Alus/g.x"
348..468
/rpc_family="FLAM_C"
469..765
/rpc_family="Alus/x"
1012..1319
/rpc_family="Alusx"
complement(1553..1867)

```

```
/rpt_family="AluJb"
complement(1868, .1977)
repeat_region
/rpt_family="AluJo"
2381, .2686
repeat_region
/rpt_family="AluSx"
2687, .2756
/rpt_family="(TA)n"
complement(2757, .2829)
repeat_region
/rpt_family="Rickenha"
complement(2830, .3124)
repeat_region
/rpt_family="AluSg"
complement(3125, .3254)
repeat_region
/rpt_family="Rickenha"
3262, .3847
repeat_region
/rpt_family="LTR9"
3966, .4259
repeat_region
/rpt_family="AluSx"
complement(4327, .4607)
repeat_region
/rpt_family="HERVL"
4833, .5124
repeat_region
/rpt_family="AluSg"
5125, .5436
repeat_region
/rpt_family="AluSx"
5635, .5683
repeat_region
/rpt_family="AT rich"
complement(5746, .5934)
repeat_region
/rpt_family="AluJb"
complement(5935, .6248)
repeat_region
/rpt_family="AluSx"
complement(6249, .6350)
repeat_region
/rpt_family="AluJb"
6613, .6747
repeat_region
/rpt_family="MER112"
complement(6837, .7158)
repeat_region
/rpt_family="AluSc"
7439, .7727
repeat_region
/rpt_family="AluSp"
7738, .7771
repeat_region
/rpt_family="(TAA)n"
complement(7828, .7948)
repeat_region
/rpt_family="MLT1H"
complement(8132, .8280)
repeat_region
/rpt_family="MLT1G"
complement(8357, .8467)
repeat_region
/rpt_family="L2"
complement(8472, .8775)
repeat_region
/rpt_family="AluSg"
complement(9131, .9257)
repeat_region
/rpt_family="L2"
9246, .9369
repeat_region
/rpt_family="MIR"
9730, .9751
repeat_region
/rpt_family="AT rich"
9761, .9893
repeat_region
/rpt_family="AluSg/x"
9907, .10200
repeat_region
/rpt_family="AluSx"
complement(10355, .10680)
repeat_region
/rpt_family="L2"
11039, .11148
repeat_region
/rpt_family="MIR"
```

Query Match 1.7%; Score 52; DB 8; Length 163157;  
Best Local Similarity 100.0%; Pred. No. 5.7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3071 CAAGATTGCGACCTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
|||||  
DB 22314 CAAGATTGCGACCTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 22263

RESULT 133  
AC141308

LOCUS AC141308 163280 bp DNA linear HTG 11-MAR-2003  
DEFINITION Homo sapiens chromosome 16 clone RP11-855H8, WORKING DRAFT  
SEQUENCE, 3 unordered pieces.  
AC141308  
AC141308.1 GI:28913088  
VERSION HTG; HTGS PHASE1; HTGS\_DRAFT; HTGS\_ACTIVEFIN.  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 163280)  
DOE Joint Genome Institute.  
TITLE Sequencing of Human Chromosome 16  
JOURNAL Unpublished  
2 (bases 1 to 163280)  
DOE Joint Genome Institute.  
AUTHORS Direct Submission  
TITLE Submitted (11-MAR-2003) Production Sequencing Facility, DOE Joint  
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94596, USA  
COMMENT  
-----Genome Center  
Center: Joint Genome Institute  
Center Code: JGI  
Web site: <http://www.jgi.doe.gov>  
-----  
Project Information  
Center Project Name: 1623000  
Center clone name: RPC1-11\_855H8  
-----

Summary Statistics  
Consensus quality: 162905 bases at least Q40  
Consensus quality: 162980 bases at least Q30  
Consensus quality: 163047 bases at least Q20  
Estimated insert size: 175000; agarose-1p estimation  
Estimated insert size: 163080; sum-of-contigs estimation  
Quality coverage: 13.59 in Q20 bases; agarose-1p estimation  
Quality coverage: 14.58 in Q20 bases; sum-of-contigs estimation.  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 3 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 22836: contig of 22836 bp in length  
\* 22837 22936: gap of unknown length  
\* 22937 49828: contig of 26892 bp in length  
\* 49829 49928: gap of unknown length  
\* 49929 163280: contig of 113352 bp in length.  
FEATURES  
source  
1..163280  
/organism="Homo sapiens"  
/mol\_type="Genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="16"  
/clone="RP11-855H8"  
/clone\_11b="RPCT human BAC library 11"  
22837..22936  
gap  
/estimated\_length=unknown  
49829..49928  
gap  
/estimated\_length=unknown

## ORIGIN

Query Match 1.7%; Score 52; DB 14; Length 163280;  
Best Local Similarity 100.0%; Pred. No. 5.7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3071 CAAGATTGCGACCTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
|||||  
DB 134135 CAAGATTGCGACCTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 134186

```

RESULT 134
AC161476      163432 bp   DNA      linear   HTG 24-JUN-2005
LOCUS         Pan trolodytes chromosome UNKNOWN clone CH251-284F7, WORKING DRAFT
DEFINITION   SEQUENCE, 16 unordered pieces.
ACCESSION    AC161476
VERSION      AC161476.1 GI:65146782
KEYWORDS     HTG: HTGS PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE       Pan trolodytes (chimpanzee)
ORGANISM     Pan trolodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Pan.

REFERENCE    1 (bases 1 to 163432)
AUTHORS      Wilson,R.K.
TITLE        The sequence of Pan trolodytes clone
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 163432)
AUTHORS      Wilson,R.K.
TITLE        Direct Submission
JOURNAL      Submitted (14-MAY-2005) Genetics, Genome Sequencing Center, 4444
              Forest Park Parkway, St. Louis, MO 63108, USA
              3 (bases 1 to 163432)
              Wilson,R.K.
              Direct Submission
              Submitted (24-JUN-2005) Genetics, Genome Sequencing Center, 4444
              Forest Park Parkway, St. Louis, MO 63108, USA

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.edu
Project name: C_AB0284F07
Center project name: C_AB0284F07
----- Summary Statistics -----
Sequencing vector: M13; 0%
Chemistry: Dye-Primer ET; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 152742 bases at least Q40
Consensus quality: 156150 bases at least Q30
Consensus quality: 158103 bases at least Q20

* NOTE: This is a 'working draft' sequence. It currently
* consists of 16 configs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the configs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1      1144: config of 1144 bp in length
*      1145      1244: gap of unknown length
*      1245      2840: config of 1596 bp in length
*      2841      2940: gap of unknown length
*      2941      5060: config of 2120 bp in length
*      5061      5160: gap of unknown length
*      5161      8540: config of 3380 bp in length
*      8541      8640: gap of unknown length
*      8641      11235: config of 2595 bp in length
*      11236      11335: gap of unknown length
*      11336      11780: config of 7445 bp in length
*      11781      18880: gap of unknown length
*      18881      24016: config of 5136 bp in length
*      24017      24116: gap of unknown length
*      24117      30109: config of 5993 bp in length
*      30110      30209: gap of unknown length
*      30210      42865: config of 12656 bp in length
*      42866      42965: gap of unknown length
*      42966      51561: config of 8596 bp in length

```

```

FEATURES
source
1..163432
/organism="Pan trolodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/chromosome="UNKNOWN"
/clone="CH251-284F7"
1..1144
/note="assembly_name:Contig41"
1145..1244
/estimated_length=unknown
1245..2840
/note="assembly_name:Contig45"
2841..2940
/estimated_length=unknown
2941..5060
/note="assembly_name:Contig50"
5061..5160
/estimated_length=unknown
5161..8540
/note="assembly_name:Contig51"
8541..8640
/estimated_length=unknown
8641..11235
/note="assembly_name:Contig52"
11236..11335
/estimated_length=unknown
11336..18780
/note="assembly_name:Contig53"
18781..18880
/estimated_length=unknown
18881..24016
/note="assembly_name:Contig54"
24017..24116
/estimated_length=unknown
24117..30109
/note="assembly_name:Contig55"
30110..30209
/estimated_length=unknown
30210..42865
/note="assembly_name:Contig56"
42866..42965
/estimated_length=unknown
42966..51561
/note="assembly_name:Contig57
vector_side:right"
clone_end:856
51562..51661
/estimated_length=unknown
51662..65150
/note="assembly_name:Contig58"
65151..65250
/estimated_length=unknown
65251..74190
/note="assembly_name:Contig59"
74191..74290
/estimated_length=unknown
74291..84891
/note="assembly_name:Contig60"
84892..84991
/estimated_length=unknown
84992..103200
/estimated_length=unknown
103201..125971
/estimated_length=unknown
125972..163432
/estimated_length=unknown

```

```

misc_feature      84992..103199
                  /note="assembly_name:Contig61"
gap              103200..103299
                  /estimated_length=unknown
misc_feature      103300..125871
                  /note="assembly_name:Contig62"
gap              125872..125971
                  /estimated_length=unknown
misc_feature      125972..163432
                  /note="assembly_name:Contig63"
ORIGIN
Query Match      1.7%; Score 52; DB 14; Length 163432;
Best Local Similarity 100.0%; Pred No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      2889 GAGGCAAGTGCATCAGCTGAGGCGAGAGTTCGAGACGAGCTGGCCAAAT 2940
          |||||
          91488 GAGGCAAGTGCATCAGCTGAGGCGAGAGTTCGAGACGAGCTGGCCAAAT 91539
RESULT 135
AC092606      163437 bp DNA linear PRI 08-OCT-2003
LOCUS      Homo sapiens BAC clone RP11-150B23 from 7, complete sequence.
DEFINITION      AC092606 AC092606
ACCESSION      AC092606.2 GI:115638770
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
              Hominoidea; Homo.
              1 (bases 1 to 163437)
              Suleston,J.E. and Wilson,R.
              Toward a complete human genome sequence
              Genome Res. 8 (11), 1097-1108 (1998)
              9847074
REFERENCE      2 (bases 1 to 163437)
AUTHORS      Abbot,A., Doeber,A. and Hou,S.
              The sequence of Homo sapiens BAC clone RP11-150B23
              Unpublished (2001)
REFERENCE      3 (bases 1 to 163437)
AUTHORS      Waterston,R.H.
              Direct Submission
              Submitted (19-JUL-2001) Genome Sequencing Center, Washington
              University School of Medicine, 4444 Forest Park Parkway, St. Louis,
              MO 63108, USA
              4 (bases 1 to 163437)
              Waterston,R.H.
              Direct Submission
              Submitted (01-MAR-2002) Department of Genetics, Washington
              University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
              MO 63108, USA
              5 (bases 1 to 163437)
              Waterston,R.
              Direct Submission
              Submitted (04-JUN-2002) Department of Genetics, Washington
              University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
              7 (bases 1 to 163437)
              Waterston,R.
              Direct Submission
              Submitted (29-APR-2003) Department of Genetics, Washington
              University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
              8 (bases 1 to 163437)
              Wilson,R.
              Direct Submission
              Submitted (08-OCT-2003) Department of Genetics, Washington

```

## COMMENT

University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
On Sep 18, 2001 this sequence version replaced gi:14916191.

-----Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu>  
Contact: [sapiens@wustl.wustl.edu](mailto:sapiens@wustl.wustl.edu)  
-----Summary Statistics  
Center project name: H\_NH0150B23  
Drafting Center: WIBR  
-----

NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.

## MAPPING INFORMATION:

The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send [mailto:egreen@nhgri.nih.gov](mailto:mailto:egreen@nhgri.nih.gov), or see <http://genome.wustl.edu>

## SOURCE INFORMATION:

The RPCI-11 human BAC library was made from the blood of one male donor, as described by Osogawa,K., Moon,P.Y., Zhao,B., Prengen,B., Tateo,M., Cacanes,J.V. and de Jong,P.V. (1998) An improved approach for construction of bacterial artificial chromosome libraries. Genomics 51:1-8. The clone may be obtained either from Research Genetics, Inc. (<http://www.regen.com>) or Pieter de Jong and coworkers at <http://www.chori.org>

VECTOR: pBAC3.6

## NEIGHBORING SEQUENCE INFORMATION:

Data from AC092148 was used to finish this clone, AC092606.

The clone sequenced to the right is RP11-95H23. Actual start of this clone is at base position 1 of RP11-150B23 actual end is at base position 163437 of RP11-150B23.

The sequence of AC0925297 has been incorporated into AC092606.

## FEATURES

```

source      Location/Qualifiers
            1..163437
            /organism="Homo sapiens"
            /mol_type="Genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="7"
            /map="7"
            /clone="RP11-150B23"
            /clone_11b="RPCI-11"
            1..953
            /rpt_family="L1"
            repeat_region
            1041..1227
            /rpt_family="L1"
            repeat_region
            1227..1868
            /rpt_family="L1"
            repeat_region
            3283..3321
            /rpt_family="L1"
            repeat_region
            10219..10601
            /rpt_family="MALR"
            repeat_region
            10602..12171

```

```
repeat_region /rpt_family="MaLR"
12172..12542
repeat_region /rpt_family="MaLR"
13005..13032
repeat_region /rpt_family="AT_rich"
13940..14015
repeat_region /rpt_family="MIR"
14231..14531
repeat_region /rpt_family="Alu"
16011..16093
repeat_region /rpt_family="MIR"
17056..17085
repeat_region /rpt_family="AT_rich"
18554..18884
repeat_region /rpt_family="Alu"
19610..19797
repeat_region /rpt_family="MIR"
19838..19908
repeat_region /rpt_family="L2"
19909..20298
repeat_region /rpt_family="MaLR"
21626..21672
repeat_region /rpt_family="L2"
22177..22416
repeat_region /rpt_family="Alu"
22180..22419
misc_feature /note="CPG_Island (%GC=58.3, o/e=0.86, #CPGs=18)"
22422..22455
repeat_region /rpt_family="(A)n"
22747..22869
repeat_region /rpt_family="MER1_type"
22871..23099
repeat_region /rpt_family="MaLR"
23115..23659
repeat_region /rpt_family="ERV1"
24116..24359
repeat_region /rpt_family="L2"
24900..24942
repeat_region /rpt_family="(CMTA)n"
25007..25236
repeat_region /rpt_family="MIR"
25878..26146
repeat_region /rpt_family="Alu"
27025..27127
repeat_region /rpt_family="L2"
27128..27167
repeat_region /rpt_family="(CA)n"
27168..27388
repeat_region /rpt_family="L2"
28936..29055
repeat_region /rpt_family="MIR"
29766..29797
repeat_region /rpt_family="(TA)n"
29878..29904
repeat_region /rpt_family="(GA)n"
29906..30195
repeat_region /rpt_family="Alu"
30620..30777
repeat_region /rpt_family="L1"
30778..31086
repeat_region /rpt_family="Alu"
31087..31289
repeat_region /rpt_family="L1"
32124..32189
repeat_region /rpt_family="L2"
32323..32442
repeat_region /rpt_family="CR1"
33086..33222
repeat_region /rpt_family="MIR"
33274..33551
repeat_region /rpt_family="Alu"
33822..34122
repeat_region /rpt_family="MaLR"
```

```
repeat_region 34435..34711
repeat_region /rpt_family="L2"
repeat_region 35362..35457
repeat_region /rpt_family="CT-rich"

Query Match 1.7% Score 52; DB 8; Length 163437;
Best Local Similarity 100.0%; Pred. No. 57e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCACATGACCTCCAGCTGAGCAACAGCAAGACTGTCTTC 3122
DB 8534 CAAGATTGTGCACATGACCTCCAGCTGAGCAACAGCAAGACTGTCTTC 85415

RESULT 136
AC096508 163662 bp DNA linear HTG 13-SEP-2002
LOCUS AC096508/C
DEFINITION Homo sapiens chromosome X clone RP11-111F17, WORKING DRAFT
ACCESSION AC096508.8 GI:22830224
VERSION HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP; HTGS_ACTIVEPIN.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
1 (bases 1 to 163662)
Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
Albrooks,S.L., Amaralunga,H.C., Are,J.R., Ayale,M., Banks,T.,
Barbata,J., Benton,J., Blinze,K., Blankenburg,K., Bonnin,D.,
Bouck,J., Bowie,S., Brileva,M., Brown,E., Brown,M., Bryant,N.P.,
Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
Carron,T.F., Carter,M., Cavazos,S.R., Chacko,V., Chavez,D.,
Chen,G., Chen,R., Chen,Z., Chowdhury,I., Christopoulos,C.,
Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
Davela,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
Delaney,K.R., Delgado,O., Dem,A.L., Ding,Y., Dint,H.,
Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
Eamhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escoto,M.,
Falls,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
Gorell,J.H., Guevara,W., Gunatane,P., Hale,S., Hamilton,K.,
Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., Hernandez,J.,
Hernandez,O., Hodgson,A., Hughes,M., Holloway,C., Hollins,B.,
Jacobson,B., Jia,Y., Johnson,R., Jolivet,S., Joudah,S.,
Karlsson,E., Kelly,S., Khan,U., Kling,L., Korvah,J., Kovar,C.,
Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C., Lewis,L.,
Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W., Louisged,H.,
Lozano,R.V., Lu,X., Lucier,A., Lucier,R., Luna,R., Ma,J.,
Maheshwari,M., Mapua,P., Martin,R., Martindale,A., Martinez,E.,
Massey,E., Mawhney,E., McLeod,M.P., Medori,M., Mei,G., Metker,M.,
Miner,G., Miner,Z., Mitchell,T., Mohabbat,K., Morgan,M., Morris,S.,
Moser,M., Neal,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N.,
Nguyen,N., Nickerson,E., Nwokkenko,S., Ogulu,M., Okunolu,G.,
Oreguine,N., Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L.,
Peters,L., Plickens,R., Primus,E., Pu,L., Quiles,M., Remy,X.,
Rives,M., Rojas,A., Rojubokan,I., Rolfe,M., Rolfe,S., Savary,G.,
Schereg,S., Scott,G., Shen,H., Shoosticari,N., Sisson,I.,
Sodergren,E., Sonalke,T., Sparks,A., Stanley,H., Stone,H.,
Sutton,A., Svalok,A., Taber,P., Tamerisa,A., Tamerisa,K., Tang,H.,
Tansey,J., Taylor,C., Taylor,T., Telford,B., Thomas,N., Thomas,S.,
Tumani,K., Vaequez,L., Verz,V., Villalon,D., Vinson,R., Wang,Q.,
Wang,S., Ward-Moore,S., Warren,R., Washington,C., Watlington,S.,
Williams,G., Williamson,A., Wleczek,K., Wooten,S., Worley,K.,
Wu,C., Wu,Y., Wu,Y.F., Zhou,J., Zorrilla,S., Nelson,D.,
Weinstock,G. and Gibbs,R.
Direct Submission
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 163662)
AUTHORS Worley,K.C.
TITLE Direct Submission
```









```

7368      7290      7551      7596      1507      1498
-----
740      <800      1891      1874      5891      5919
-----
6924      6877      3698      3748      8949      9128
-----
488      <800      218      <800      2678      2827
-----
1689      1615      11309      11212      51      <800
-----
634      <800      770      <800      6768      6510
-----
376      <800      8011      8134      3884      3943
-----
12088     12061     581      <800      2500      2661
-----
-----
1780      1874      4918      4800
-----
641      <800      719      <800
-----
5216      5119      2412      2419
-----
2255      2322
-----
1179      1120
-----
734      <800
-----

FEATURES
source
1..165311
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-415L24"
/clone_lib="RP11 human BAC library 11"

ORIGIN

Query Match      1.7% Score 52; DB 8; Length 165311;
Best Local Similarity 100.0%; Pred.No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      3071 CAAGATTGCGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122
      |||||
Db      60695 CAAGATTGCGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 60644

RESULT 139
AC135854      165972 bp      DNA      linear      HTG 27-MAR-2003
LOCUS      AC135854/c
DEFINITION      Homo sapiens chromosome 15 clone RP11-1404H5 map 15, 3 unordered
pieces.
AC135854
AC135854.4      GI:27552578
HTG; HTGS PHASE1; HTGS_FUL1TOP; HTGS_CANCELLED.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 165972)
Birren,B., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 15, clone RP11-1404H5
Unpublished
2 (bases 1 to 165972)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barra,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
Cammarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,

```

```

Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J.,
Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,U., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-OCT-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 165972)
REFERENCE
JOURNAL
AUTHORS
Birren,B., Nusbaum,C., Lander,B., Ali,A., Allen,N., Anderson,S.,
Barra,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhgalter,B.,
Cammarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J.,
Matthews,C., McCarthy,M., Meldrim,J., Meneus,L., Mihova,T.,
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,
Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K.,
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,U., Zimmer,A. and Zody,M.
Direct Submission
Submitted (27-MAR-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jan 9, 2003 this sequence version replaced gi:27531844.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
-----
Project Information
Center project name: 1404_H_5
Center clone name: 1404_H_5
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 3 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 22900: contig of 22900 bp in length
* 22901 23000: gap of 100 bp
* 23001 138676: contig of 115676 bp in length
* 138677 138776: gap of 100 bp
* 138777 165972: contig of 27196 bp in length.
-----
FEATURES
source
1..165972
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-1404H5"
/clone_lib="RP11 human BAC segment 5"
22901..23000
22901..23000
138677..138776
138677..138776

```

ORIGIN /estimated\_length=100

Query Match 1.7%; Score 52; DB 14; Length 165972;  
Best Local Similarity 100.0%; Pred. No. 5.7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2889 GAGCAGGTGATCAGCTGAGGCGCAGAGTTCAGACGACCTGGCCACAT 2940  
|||||  
Db 76245 GAGCAGGTGATCAGCTGAGGCGCAGAGTTCAGACGACCTGGCCACAT 76194

## RESULT 140

AC026830 166484 bp DNA linear HTG 24-AUG-2002  
LOCUS AC026830  
DEFINITION Homo sapiens chromosome 15 clone RP11-215H14 map 15, WORKING DRAFT  
SEQUENCE, 32 unordered pieces.  
AC026830  
AC026830.2 GI:7596890  
HTG: HTGS PHASE1; HTGS\_DRAFT.  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM

REFERENCE  
AUTHORS Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
TITLE Homo sapiens chromosome 15, clone RP11-215H14  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 166484)

Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
Anderson, S., Baldwin, J., Barna, N., Baetien, V., Bede, F.,  
Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G.,  
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,  
Collimore, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S.,  
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howard, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J.,  
Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N.,  
McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McSheeters, R.,  
Meldrum, J., Menues, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,  
Murphy, T., Naylor, T.M., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., P.,  
Pisenti, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Kochman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
Testafte, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,  
Vasilev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (24-MAR-2000) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 166484)

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
Anderson, S., Baldwin, J., Barna, N., Baetien, V., Bede, F.,  
Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G.,  
Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,  
Collimore, A., Cooke, P., DeArrellano, K., Dewar, K., Diaz, J.S.,  
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,  
Howard, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
Klein, J., Lacroque, K., Lamazares, R., Landers, T., Lehoczy, J.,  
Levine, R., Lieu, C., Liu, G., Locke, K., MacDonald, P., Marquis, N.,  
McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McSheeters, R.,  
Meldrum, J., Menues, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J.,  
Murphy, T., Naylor, T.M., Norman, C.H., O'Connor, T., O'Donnell, P.,  
O'Neill, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N., P.,  
Pisenti, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Kochman, D.,  
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,

TITLE  
JOURNAL  
COMMENT

Testafte, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,  
Vasilev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
Young, G., Zainoun, J., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Apr 19, 2000 this sequence version replaced gi:7321633.  
All repeats were identified using RepeatMasker:  
Smit, A.P.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIRB

Web site: http://www-seq.wi.mit.edu

Contact: sequence\_submissions@genome.wi.mit.edu

Project information

Center clone name: 215\_H\_14

Center project name: L7706

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 150453 bases at least Q40

Consensus quality: 158028 bases at least Q30

Consensus quality: 161192 bases at least Q20

Insert size: 176000; agarose-fp

Insert size: 163384; sum-of-contigs

Quality coverage: 3.6 in Q20 bases; agarose-fp

Quality coverage: 3.9 in Q20 bases; sum-of-contigs

Quality coverage: 3.9 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 32 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1025 1024: contig of 1024 bp in length  
\* 1025 1124: gap of 100 bp  
\* 1125 2485: contig of 1361 bp in length  
\* 2486 2485: gap of 100 bp  
\* 2586 3626: contig of 1041 bp in length  
\* 3627 3726: gap of 100 bp  
\* 3727 4884: contig of 1158 bp in length  
\* 4885 4985: gap of 100 bp  
\* 4986 6513: contig of 1529 bp in length  
\* 6514 6513: gap of 100 bp  
\* 6515 8513: contig of 1900 bp in length  
\* 8514 8513: gap of 100 bp  
\* 8515 10616: contig of 2003 bp in length  
\* 10617 10716: gap of 100 bp  
\* 10717 12445: contig of 1729 bp in length  
\* 12446 12545: gap of 100 bp  
\* 12546 16043: contig of 3498 bp in length  
\* 16044 16143: gap of 100 bp  
\* 16144 19800: contig of 3657 bp in length  
\* 19801 19800: gap of 100 bp  
\* 19802 23755: contig of 3855 bp in length  
\* 23756 23855: gap of 100 bp  
\* 23856 25033: contig of 1178 bp in length  
\* 25034 25133: gap of 100 bp  
\* 25134 28296: contig of 3163 bp in length  
\* 28297 31054: gap of 100 bp  
\* 31055 31154: gap of 100 bp  
\* 31155 35420: contig of 4266 bp in length  
\* 35421 35520: gap of 100 bp  
\* 35521 39320: contig of 3700 bp in length  
\* 39321 39320: gap of 100 bp  
\* 39322 42676: contig of 3356 bp in length  
\* 42677 42776: gap of 100 bp  
\* 42777 46418: contig of 3642 bp in length

```

* 46519 46518: gap of 100 bp
* 46519 51021: contig of 4503 bp in length
* 51022 51121: gap of 100 bp
* 51122 56073: contig of 4952 bp in length
* 56074 56173: gap of 100 bp
* 56174 61048: contig of 4875 bp in length
* 61049 61148: gap of 100 bp
* 61149 66205: contig of 5057 bp in length
* 66206 66305: gap of 100 bp
* 66306 72086: contig of 5781 bp in length
* 72087 72187: gap of 100 bp
* 72188 77190: contig of 5004 bp in length
* 77191 77290: gap of 100 bp
* 77291 87414: contig of 10124 bp in length
* 87415 87514: gap of 100 bp
* 87515 94355: contig of 6841 bp in length
* 94356 103495: gap of 100 bp
* 103496 103595: contig of 9040 bp in length
* 103596 112676: gap of 100 bp
* 112677 112776: contig of 9081 bp in length
* 112777 122142: gap of 100 bp
* 122143 122242: contig of 9366 bp in length
* 122243 131715: contig of 9473 bp in length
* 131716 131815: gap of 100 bp
* 131816 142646: contig of 10831 bp in length
* 142647 142747: gap of 100 bp
* 142748 166484: contig of 23738 bp in length.
Location/Qualifiers
source
1..166484
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="15"
/map="15"
/clone="RP11-215H14"
/clone_lib="RPC1-11 Human Male BAC"
1..1024
/note="assembly_fragment"
gap
1025..1124
/estimated_length=100
misc_feature
1125..2485
/note="assembly_fragment"
2486..2585
/estimated_length=100
misc_feature
2586..3626
/note="assembly_fragment"
3627..3726
/estimated_length=100
misc_feature
3727..4884
/note="assembly_fragment"
4885..4984
/estimated_length=100
misc_feature
4985..6513
/note="assembly_fragment"
6514..6613
/estimated_length=100
misc_feature
6614..8513
/note="assembly_fragment"
8514..8613
/estimated_length=100
misc_feature
8614..10616
/note="assembly_fragment"
10617..10716
/estimated_length=100
misc_feature
10717..12445
/note="assembly_fragment"
12446..12545

```

```

Query Match 1.7%; Score 52; DB 14; Length 166484;
Best Local Similarity 100.0%; Pred. No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2889 GAGGAGGTGATCATCCTGAGGCCAGAGTTCGAGACCAAGCCTGGCAACAT 2940
Db 67810 GAGGAGGTGATCATCCTGAGGCCAGAGTTCGAGACCAAGCCTGGCAACAT 67861

RESULT 141
LOCUS AL844892
DEFINITION Human DNA sequence from clone RP11-396M20 on chromosome 10 Contains
AL844892 the 5' end of the GRID1 gene for glutamate receptor (ionotropic)
delta 1, the 3' end of the gene for friend of EBNM2 (FOS)
(K18A0261) and five CpG islands, complete sequence.
ACCESSION AL844892.5 GI:23395595
VERSION HTG; FOS; GRID1; K18A0261.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo
1 (bases 1 to 167780)
REFERENCE 1
AUTHORS Howden,P.
TITLE Direct Submission
JOURNAL Submitted (17-MAY-2005) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegase@sanger.ac.uk
COMMENT On Oct 1, 2002 this sequence version replaced gi:22798583.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 10, constructed by the Sanger Centre Chromosome 10
Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr10
RP11-396M20 is from the library RPC1-11.2 constructed by the group
of Pletzer de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pBAC3.6
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vegase@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
Location/Qualifiers
source
1..167780
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="10"
/clone="RP11-396M20"
/clone_lib="RPC1-11.2"
1
/note="Clone left end: RP11-396M20"
48351..48357
/note="Sequence from uni-directional dGTP big dye
terminator reads only."
57532
/note="Clone right end: RP11-539A10"
join(complement(AL683834.9:36029..103858),
complement(AC073162.9:114306..1176049),
complement(AL596135.6:7440..14051),
complement(AC022028.7:40751..86154))

```

```

/gene="GRID1"
/locus_tag="RP11-93H12.1-001"
join(complement(61169..61348),complement(58796..58951),
complement(AL683834.9:103574..103858),
complement(AL683834.9:136029..13634),
complement(AC073162.9:175996..176049),
complement(AC073162.9:128820..128990),
complement(AC073162.9:115839..116000),
complement(AC073162.9:114306..114425),
complement(AC073162.9:113950..14051),
complement(AL596135.6:13292..12489),
complement(AL596135.6:8789..9113),
complement(AL596135.6:7440..7578),
complement(AC022028.7:85959..86154),
complement(AC022028.7:52164..52404),
complement(AC022028.7:40751..41458))
/gene="GRID1"
/locus_tag="RP11-93H12.1-001"
/product="glutamate receptor, ionotropic, delta 1"
/notes="match: cDNAs: BC039263.1"
join(complement(61169..61247),complement(58796..58951),
complement(AL683834.9:103574..103858),
complement(AL683834.9:136029..136234),
complement(AC073162.9:175996..176049),
complement(AC073162.9:128820..128990),
complement(AC073162.9:115839..116000),
complement(AC073162.9:114306..114425),
complement(AL596135.6:13950..14051),
complement(AL596135.6:12292..12489),
complement(AL596135.6:8789..9113),
complement(AL596135.6:7440..7578),
complement(AC022028.7:85959..86154),
complement(AC022028.7:52350..52404))
/gene="GRID1"
/locus_tag="RP11-93H12.1-001"
/standard_name="OTTHUMP0000020002"
/notes="match: proteins: Q61627 Q62640"
/codon_start=1
/product="glutamate receptor, ionotropic, delta 1"
/protein_id="CAH73378.1"
/db_xref="GI:5566097"
/db_xref="GOA:Q5VSK3"
/db_xref="InterPro:IPR001311"
/db_xref="InterPro:IPR001320"
/db_xref="InterPro:IPR001508"
/db_xref="InterPro:IPR001828"
/db_xref="UniProt/TREMBL:Q5VSK3"
/translacion="MEALTLMLPWCQCVCVRAVSIHIGAFIENAKXDRVFLA
VDSLISLNDILOSEKITYSIKYIEANNPFOAEOACDLMTQGLALVSTGASNAL
OSLTAMH1PHL.FVORNPQSPGSPRTCHLNPSPGKAYTLASRPVRINDVMLRIATEL
RROKTVFYDSYDRIGLOSPLDQASRLDVLSDQVKNISHVFTSLFTTKTEBN
RTRDLTKRAILILSPQASHFTINEAETNLAKDSWVNSEISDPELIXDVLHSLG
HMTVVRQIIPPSADNOKCTRNHRISSLLCDQEGTQMLQSNLYLXSDVLMANF
RKTLBDRKWHNSASNLICIRKSTKPMNGGSMIDTIKGIHITGLTGVMEBRESNPYV
OEILGITYSETFGDKMRKLTWDSKGLNGSLOSPRGSRLQGLIKAVTLLEBPV
WAENLIGOPKRYKGSIDVLDALAKALGFKYKTYQAPRGYGHOLAHNSNMGCEL
ISKRADLAISATITPERESVDFSKRTYDVSGLILIKPEEKISIFSLFAPFDRAVW
ACTIAAIVPGVGLFVLANRIOAVRAQSAOPPSASATLHSAIWIYVGAFOQGSSESS
VNSAMRIYVGSWMLFTLVCSYSYANLAAPLTVSMNPRIPTFDLSKQVMSVGTV
RDSAYVEYFRAGKTGPLEBQDSTFAELMRTISKNGADNCVSSPSEIGIRDPEAGHRG
PCCAEAHVVAHGPL"
61673..61783
/notes="Single clone region. Sequence generated from a
transposon library derived from a single pUC clone.
Restriction digest data confirm the assembly."
join(complement(AL731569.18:50750..50890),
complement(AL731569.18:50440..50616),
complement(AL731569.18:46646..47166),
complement(AL731569.18:28793..29818),
complement(AL731569.18:26276..26394),
complement(AL731569.18:2953..3048),
complement(167467..167619),complement(167063..167199),
complement(167467..167619),complement(162134..162361),
complement(165847..165958),complement(162134..162361),

```

```

complement(156031..156142),complement(155254..155351),
complement(153844..153920),complement(148511..148686),
complement(148075..148190),complement(146822..146986),
complement(141097..141304),complement(138130..138218),
complement(132775..132870),complement(130111..132463))
/gene="XIAA0261"
/locus_tag="RP11-396M20.1-001"
join(complement(AL731569.18:50750..50890),
complement(AL731569.18:50440..50616),
complement(AL731569.18:46646..47166),
complement(AL731569.18:28793..29818),
complement(AL731569.18:26276..26394),
complement(AL731569.18:2953..3048),
complement(167467..167619),complement(167063..167199),
complement(165847..165958),complement(162134..162361),
complement(156031..156142),complement(155254..155351),
complement(153844..153920),complement(148511..148686),
complement(148075..148190),complement(146822..146986),
complement(141097..141304),complement(138130..138218),
complement(132775..132870),complement(130111..132463))
/gene="XIAA0261"
/locus_tag="RP11-396M20.1-001"
/product="XIAA0261"
/notes="match: ESTs: AA579854.1 AW577904.1 B1438658.1
B1459087.1 BM504319.1 BM683407.1 BQ007553.1 BQ420406.1
BU620920.1
match: cDNAs: D87450.1"
complement(join(130111..132463,132775..132870),
138130..138887))
/gene="XIAA0261"
/locus_tag="RP11-396M20.1-002"
complement(join(130111..132463,132775..132870),
138130..138887))
/gene="XIAA0261"
/locus_tag="RP11-396M20.1-002"
/product="XIAA0261"
/notes="match: cDNAs: BC036507.1"
complement(130111)
/gene="XIAA0261"
/locus_tag="RP11-396M20.1-002"
complement(130129..130134)
/gene="XIAA0261"
/locus_tag="RP11-396M20.1-002"
complement(join(131865..133463,133775..132870),
138130..138218,141097..141304,146075..148190,
148511..148686,153844..153920,155254..155351,
156031..156142,162134..162363))
/gene="XIAA0261"
/locus_tag="RP11-396M20.1-004"
complement(join(131865..133463,133775..132870),
138130..138218,141097..141304,146075..148190,
148511..148686,153844..153920,155254..155351,
156031..156142,162134..162363))
/gene="XIAA0261"
/locus_tag="RP11-396M20.1-004"
/product="XIAA0261"
/notes="match: cDNAs: BC017393.1"
join(complement(AL731569.18:50750..50890),
complement(AL731569.18:50440..50616),
complement(AL731569.18:46646..47166),
complement(AL731569.18:28793..29818),
complement(167467..167619),complement(167063..167199),
complement(165847..165958),complement(162134..162361),
complement(156031..156142),complement(155254..155351),
complement(153844..153920),complement(148511..148686),
complement(148075..148190),complement(146822..146986),
complement(141097..141304),complement(138130..138218),
complement(132775..132870),complement(130111..132463))

```

QY	3071	CAAGATTGTCACGACGACCTGCAGGCTGGGACAGAGAGACGTCGTCTC	3122
Db	93241	CAAGATTGTCACGACGACCTGCAGGCTGGGACAGAGAGACGACGTCGTCTC	93292

RESULT 142  
ACT109635

Query Match 1.7%; Score 52; DB 8; Length 167780;  
Best Local Similarity 100.0%; Pred. No. 5.7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;



LOCUS	AC109635	167920 bp	DNA	linear	PRI 10-SEP-2002
DEFINITION	Homo sapiens chromosome 11, clone RP11-347H15, complete sequence.				
ACCESSION	AC109635				
VERSION	AC109635.6	GI:227273373			
KEYWORDS	HTG.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.				
REFERENCE	1 (bases 1 to 167920) Birren, B., Nusbaum, C. and Lander, E. Homo sapiens chromosome 11, clone RP11-347H15 Unpublished				
AUTHORS	2 (bases 1 to 167920) Birren, B., Linton, L., Nusbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Batta, N., Baetien, V., Boguslavskiy, L., Boukhalter, B., Brown, A., Camarata, J., Campopiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fato, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S., Gilde, S., Gord, S., Goyette, M., Graham, L., Grand-pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Lander, C., Laroque, K., Lamazza, R., Landers, I., Lehotsky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Margus, N., Matthews, C., McCarthy, M., Mcwan, P., McKernan, K., Meldrim, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicoll, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Strauss, N., Subramanian, A., Talamas, J., Teefaye, S., Theodore, J., Topham, K., Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.				
TITLE	Direct Submission				
JOURNAL	Submitted (06-FEB-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA				
REFERENCE	3 (bases 1 to 167920)				
AUTHORS	Birren, B., Nusbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Batta, N., Baetien, V., Bloom, T., Boguslavskiy, L., Boukhalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fato, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, I., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicoll, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Teefaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zemdek, L., Zimmer, A. and Zody, M.				
TITLE	Direct Submission				
JOURNAL	Submitted (25-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA				
REFERENCE	4 (bases 1 to 167920)				
AUTHORS	Birren, B., Nusbaum, C., Lander, E., All, A., Allen, N., Anderson, S., Batta, N., Baetien, V., Bloom, T., Boguslavskiy, L., Boukhalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Fato, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, I., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mlenga, V.,				

TITLE	Murphy, T., Naylor, J., Nguyen, C., Niccol, R., Norbu, C., Norman, C., H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Petersen, K., Phunhkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuppback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange, T., Thamm, N., Stojanovic, N., Talamas, J., Teejaye, S., Theodore, J., Toplam, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zaimoun, J., Zembek, L., Zimmerman, A. and Zody, M.
JOURNAL	Direct Submissions
COMMENT	Submitted (10-SEP-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA On Sep 10, 2002 this sequence version replaced gt:22475361. All repeats were identified using RepeatMasker: Smt, A.F.A. & Green, P. (1996-1997) http://ftp.genome.washington.edu/RM/RepeatMasker.html ----- Genome Center Center: Whitehead Institute/ MIT Center for Genome Research Center code: WtBR Web site: http://www-seq.wi.mit.edu Contact: sequence_submissions@genome.wi.mit.edu ----- Project Information Center project name: L24041 Center clone name: 347_H_15 -----
FEATURES	location/Qualifiers 1..167920 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="11" /map="11" /clone="RP11-347H15" /clone_1db="RP11-11 Human Male BAC" complement(903..2453) /rpt_family="PTR5" 2562..2726 /rpt_family="MIR" /rpt_family="MIR" complement(3163..3289) /rpt_family="AluDb" 3377..3426 /rpt_family="CATRA n" 3701..5177 /rpt_family="THE1-int" 5409..5446 /rpt_family="CA n" 6146..6179 /rpt_family="TTTG n" complement(6551..6861) /rpt_family="AluSx" 6868..6965 /rpt_family="MIR" 9236..9316 /rpt_family="L2" 9478..9544 /rpt_family="TCRA n" complement(10278..10409) /rpt_family="FLAM_C" 10439..10729 /rpt_family="AluY" complement(10738..10947) /rpt_family="L1ME" complement(10968..11095) /rpt_family="L1ME4A" complement(11142..11423) /rpt_family="L1ME4A" 11462..11535 /rpt_family="TATATG n" 11813..11846 /rpt_family="AT_rich" complement(13337..13345) /rpt_family="PRAM/FLM" 13679..13724 /rpt_family="TATATG n"

```

repeat_region      13725..13893
                    /rpt_family="(TA)n"
repeat_region      13896..14082
                    /rpt_family="(TA)n"
repeat_region      14083..14248
                    /rpt_family="(TA)n"
repeat_region      14272..14947
                    /rpt_family="(TA)n"
repeat_region      14990..15155
                    /rpt_family="(TA)n"
repeat_region      15202..15371
                    /rpt_family="(TA)n"
repeat_region      15378..15398
                    /rpt_family="AT_rich"
repeat_region      15411..15730
                    /rpt_family="(TA)n"
repeat_region      15735..15761
                    /rpt_family="AT_rich"
repeat_region      15766..15926
                    /rpt_family="(TA)n"
repeat_region      15942..16119
                    /rpt_family="(TA)n"
unsure             16041..16047
                    /note="<30 qual SNGI region"
repeat_region      complement(16537..16909)
                    /rpt_family="L2"
repeat_region      16930..16965
                    /rpt_family="(TC)n"
repeat_region      17126..17420
                    /rpt_family="AluSx"
repeat_region      complement(17423..17566)
                    /rpt_family="L1PA2"
repeat_region      17567..18711
                    /rpt_family="L1PA2"
repeat_region      complement(19966..20063)
                    /rpt_family="MIR"
repeat_region      complement(20925..20973)
                    /rpt_family="L1MC4"
repeat_region      20977..22553
                    /rpt_family="PTR5"
repeat_region      complement(22640..23478)
                    /rpt_family="L1MC4"
repeat_region      23560..23902
                    /rpt_family="L12"
repeat_region      complement(23943..24673)
                    /rpt_family="L1R8"
repeat_region      24699..24746
                    /rpt_family="(CGTG)n"
repeat_region      25464..25498
                    /rpt_family="MADEL"
repeat_region      26406..26725
                    /rpt_family="L1PA7"
repeat_region      26726..26765
                    /rpt_family="AT_rich"

Query Match      1.7% Score 52; DB 8; Length 167920;
Best Local Similarity 100.0%; Pred.No.5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      3071 CAGATTGTGCACCTGCCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122
Db      57209 CAAAGTTGTGCACCTGCCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 57260

-RESULT 143
AC007371      AC007371      168734 bp      DNA      linear      PRI 16-SEP-1999
DEFINITION      Homo sapiens, clone 24_A_9, complete sequence.
ACCESSION      AC007371
VERSION      AC007371.16      GI:5542041
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens

```

```

REFERENCE
AUTHORS      Birren,B., Linton,L., Nusbaum,C. and Lander,B.
TITLE      1 (bases 1 to 168734)
JOURNAL      Homo sapiens, clone 24_A_9
REFERENCE
AUTHORS      2 (bases 1 to 168734)
TITLE      Direct Submission
JOURNAL      Submitted (24-APR-1999) Whitehead Institute/MIT Center for Genome
REFERENCE
AUTHORS      Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 168734)
Birren,B., Linton,L., Nusbaum,C., Lander,B., Allen,N., Anderson,M.,
Baladin,J., Barna,N., Boguslavsky,L., Boukhvalter,B., Brown,A.,
Colangelo,M., Collins,S., Collum,A., Cooke,P., Dearellano,K.,
Dewar,K., Domingo,M., Donelan,L., Doyle,M., Ferreira,P.,
Fitzhugh,W., Funke,R., Gage,D., Galagan,J., Gardyna,S., Grant,G.,
Hagos,B., Heathford,A., Horton,L., Howland,J.C., Jones,C., Kann,L.,
Karatas,A., Klein,J., Lehoczy,J., Liu,C., Locke,K., MacDonald,P.,
Marquis,N., McEwan,P., McGurk,A., McKernan,K., McLaughlin,J.,
Meldrum,J., Morrow,J., Naylor,J., Norman,C.H., O'Donnell,P.,
Peterson,K., Pollara,V., Riley,R., Roy,A., Sanches,R., Severy,P.,
Strange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Testaye,S., Tirrell,A., Vasilleev,H., Vo,A., Wheeler,J., Wu,X.,
Wyman,D., Ye,W.J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (16-SEP-1999) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 21, 1999 this sequence version replaced gi:5468556.
All repeats were identified using RepeatMasker: Smit, A.P.A. &
Green, P. (1996-1997)
http://ftp.genome.washington.edu/RW/RepeatMasker.html.
FEATURES
source      1..168734
            /organism="Homo sapiens"
            /mol_type="genomic DNA"
            /db_xref="taxon:9606"
            /clone="24_A_9"
            /clone_lib="Alan Buckler -- per comm"
            complement(1444..1552)
            /rpt_family="MIR"
            complement(1581..1792)
            /rpt_family="MER20"
            complement(3654..3953)
            /rpt_family="AluSg"
            4172..4695
            /rpt_family="AluSg"
            repeat_region      4700..4730
            /rpt_family="L1MC4"
            4700..4730
            /rpt_family="(TTTG)n"
            complement(4731..5015)
            /rpt_family="AluSg"
            5016..5038
            /rpt_family="L1MC4"
            complement(5039..5331)
            /rpt_family="AluY"
            5332..5417
            /rpt_family="L1MC4"
            5490..5649
            repeat_region

```

```

repeat_region /rpt_family="MIR"
6091. .6249
/rpt_family="MIR"
repeat_region /rpt_family="MIR"
6615. .6645
/rpt_family="AT_rich"
6920. .7228
/rpt_family="AluSg"
complement(7875. .8062)
/rpt_family="L2"
repeat_region /rpt_family="MIR"
8069. .8161
complement(10637. .10816)
/rpt_family="L1MC4"
repeat_region complement(11031. 11134)
/rpt_family="MER112"
11375. .11507
/rpt_family="FLAM_C"
11901. .11932
/rpt_family="AT_rich"
12275. .12296
/rpt_family="AT_rich"
12424. .12632
/rpt_family="MIR"
12877. .13606
/rpt_family="L2"
14281. .14308
/rpt_family="AT_rich"
14414. .15069
/rpt_family="L2"
15459. .15589
/rpt_family="FLAM_C"
15650. .16082
/rpt_family="L2"
repeat_region /rpt_family="L2"
16280. .16311
/rpt_family="(TTTG)n"
complement(16862. .17059)
/rpt_family="MIR"
17113. .17366
/rpt_family="MIR"
17405. .17439
/rpt_family="(TG)n"
17591. .17645
/rpt_family="(CA)n"
complement(18182. .18482)
/rpt_family="AluSx"
complement(19280. .19588)
/rpt_family="AluDb"
19617. .19903
/rpt_family="AluDo"
19904. .19934
/rpt_family="(CAA)n"
20626. .20927
/rpt_family="AluDb"
21346. .21380
/rpt_family="AT_rich"
22056. .22353
/rpt_family="AluSg"
22650. .22705
/rpt_family="L2"
repeat_region complement(22705. .22823)
/rpt_family="MIR"
24066. .24223
/rpt_family="MLT1J2"
24294. .24435
/rpt_family="MLT1J1"
complement(24436. .24606)
/rpt_family="AluSg/x"
24646. .24717
/rpt_family="L2"
24754. .24994
/rpt_family="MLT1B"
25223. .25277
/rpt_family="MIR"

```

```

repeat_region complement(25440. .25546)
/rpt_family="MIR"
repeat_region 25616. .25690
/rpt_family="MER103"
repeat_region 25885. .26015
/rpt_family="AluSg/x"
complement(26034. .26094)
/rpt_family="MIR"
26128. .26303
/rpt_family="MIR"
repeat_region /rpt_family="MIR"
26704. .26960
/rpt_family="AluSg1"
complement(27595. .27869)
/rpt_family="AluSx"
28051. .28369
/rpt_family="AluSx"
28869. .28914
/rpt_family="AT_rich"
29087. .29157
/rpt_family="MER47"
30863. .31007
/rpt_family="MIR"
complement(31619. .31897)
/rpt_family="AluDb"
complement(31920. .32222)
/rpt_family="AluSg1"
complement(32390. .32635)
/rpt_family="L2"
/rpt_family="L2"
complement(33248. .33322)
/rpt_family="L2"
complement(33330. .33542)
/rpt_family="MIR"
33767. .33782
/rpt_family="AT_rich"
35250. .35296
/rpt_family="AT_rich"
35406. .35454
/rpt_family="(TG)n"
complement(35466. .35758)
/rpt_family="AluSx"
38644. .38729
/rpt_family="L2"
38842. .38868
/rpt_family="(TTTC)n"
38873. .38937
/rpt_family="L2"
39359. .39411
repeat_region

```

Query Match 1.7%; Score 52; DB 8; Length 168734;  
 Best Local Similarity 100.0%; Pred.No. 5.7e-16;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTCACGTGCACTCCAGCCTTGCGCAACAGACAGACTCTGTCTC 3122  
 Db 28288 CAAGATTGTCACGTGCACTCCAGCCTTGCGCAACAGACAGACTCTGTCTC 28339

## RESULT 144

AC025787 AC025787 168793 bp DNA linear PRI 20-FEB-2002  
 LOCUS Homo sapiens chromosome 11, clone RP11-227P3, complete sequence.  
 AC025787  
 AC025787  
 AC025787.9 GI:18699972  
 KEYWORDS HTG.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homiidae; Homo  
 1 (bases 1 to 168793)  
 REFERENCES  
 AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.  
 TITLE Homo sapiens chromosome 11, clone RP11-227P3  
 JOURNAL Unpublished

REFERENCE  
AUTHORS  
2 (bases 1 to 168793)  
Birtten,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,  
Anderson,S., Baldin,J., Barna,N., Bastien,V., Bede,F.,  
Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G.,  
Campopiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,  
Collamore,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S.,  
Dodge,S., Domino,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,  
Galagan,J., Gardyna,S., Ginde,S., Goyette,M., Graham,L.,  
Grand-Pierre,N., Grant,G., Hages,B., Heatford,A., Horton,L.,  
Howard,J.C., Iliev,I., Johnson,R., Jones,C., Kann,L., Karatas,A.,  
Klein,U., Lacroque,K., Lamazares,R., Landers,T., Lehoczy,J.,  
Levine,R., Lieu,C., Liu,G., Locke,K., Macdonald,P., Margulis,N.,  
McCarthy,M., McEwan,P., McGuirk,A., McKernan,K., McHeaters,R.,  
Melrim,J., Meneus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,  
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,  
O'Neill,D., Olivari,T.M., Oliver,J., Peterson,K., Pierre,N.,  
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,  
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,  
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,  
Testaye,S., Theodore,J., Tirrell,A., Travers,M., Trigilio,J.,  
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,  
Young,G., Zainoun,J., Zimmer,A. and Zody,M.

TITLE  
JOURNAL  
Submitted (14-MAR-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 168793)

REFERENCE  
AUTHORS  
Birtten,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,  
Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B.,  
Brown,A., Camarata,J., Campopiano,A., Chang,J., Chazaro,B.,  
Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A.,  
Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S.,  
Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S.,  
Ginde,S., Gird,S., Goyette,M., Graham,L., Grand-Pierre,N.,  
Hages,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
Kann,L., Karatas,A., Kelle,C., Lacroque,K., Lamazares,R.,  
Landers,T., Lehoczy,J., Levine,R., Liu,G., Maclean,C.,  
Macdonald,P., Major,J., Margulis,N., Matthews,C., McCarthy,M.,  
McEwan,P., McKernan,K., Meldrum,J., Meneus,L., Mihova,T.,  
Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,  
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,  
Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C.,  
Retra,R., Rieback,M., Riley,R., Rise,C., Rogov,P., Roman,J.,  
Rosetti,M., Roy,A., Santos,R., Schauer,S., Schnpbach,R., Seaman,S.,  
Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N.,  
Struss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J.,  
Topham,K., Travers,M., Travls,N., Trigilio,J., Vassiliev,H.,  
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.,  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE  
JOURNAL  
Submitted (18-FEB-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 168793)

TITLE  
JOURNAL  
COMMENT  
Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.  
Direct Submission  
Submitted (20-FEB-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Feb 18, 2002 this sequence version replaced gi:18377313.  
All repeats were identified using RepeatMasker:  
Smit, A.P.A. & Green, P. (1996-1997)  
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WISR  
Web site: <http://www-seq.wi.mit.edu>  
Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
----- Project Information  
Center project name: 16819  
Center clone name: 227\_P\_3  
-----  
FEATURES  
source  
Location/Qualifiers  
1..168793  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11"  
/clone="RP11-227P3"  
/clone\_1id="RP11-1 Human Male BAC"  
complement(1..97)  
/rpt\_family="L2"  
100..393  
/rpt\_family="AluDb"  
complement(648..1118)  
/rpt\_family="MER1A"  
complement(1420..1573)  
/rpt\_family="LTR41"  
1574..1717  
/rpt\_family="AluDb"  
1718..1757  
/rpt\_family="AluDb"  
complement(1758..1925)  
/rpt\_family="AluDb"  
2227..2530  
/rpt\_family="AluSg"  
complement(12540..2677)  
/rpt\_family="MIR"  
3942..3968  
/rpt\_family="AT-rich"  
4486..4792  
/rpt\_family="AluSg"  
complement(5142..5451)  
/rpt\_family="AluSg"  
complement(5848..6129)  
/rpt\_family="AluY"  
complement(6442..6567)  
/rpt\_family="MER117"  
7601..7893  
/rpt\_family="AluSg"  
8001..8296  
/rpt\_family="AluSg"  
8385..8663  
/rpt\_family="AluY"  
complement(8640..8645)  
/note="30 qual SINGL region"  
complement(8840..9137)  
/rpt\_family="AluSg"  
complement(9183..9481)  
/rpt\_family="AluSg"  
9630..9754  
/rpt\_family="AT-rich"  
9861..9889  
/rpt\_family="AluDb"  
complement(9894..10195)  
/rpt\_family="AluSg"  
10196..10467

repeat_region	/rpt_family="AluJo"	10876. .11186
unsure	/rpt_family="L2"	11226. .11229
repeat_region	/note="Probably TTTT; possibly TTT"	complement(11286. .11583)
repeat_region	/rpt_family="AluSg"	11621. .11699
repeat_region	/rpt_family="L2"	complement(11700. .11920)
repeat_region	/rpt_family="MER58A"	11921. .12456
repeat_region	/rpt_family="L2"	12523. .12649
repeat_region	/rpt_family="AluJb"	12652. .12866
repeat_region	/rpt_family="AluJo"	12870. .13185
repeat_region	/rpt_family="AluXx"	13191. .13308
repeat_region	/rpt_family="L2"	complement(13535. .13839)
repeat_region	/rpt_family="AluSg"	15032. .15436
repeat_region	/rpt_family="L1MCS"	15459. .15550
repeat_region	/rpt_family="L1MCS"	15597. .15669
repeat_region	/rpt_family="AluSg"	complement(15870. .20120)
repeat_region	/rpt_family="AluR6/alpha"	20121. .20182
repeat_region	/rpt_family="(TA)n"	complement(20183. .21439)
repeat_region	/rpt_family="AluR6/alpha"	21966. .22132
repeat_region	/rpt_family="MIR3"	complement(22855. .23061)
repeat_region	/rpt_family="L2"	complement(22904. .22918)

Query Match	Similarity	1.7%	Score 52	DB 8	Length 168793
Best Local	Similarity	100.0%	Pred. No. 5,7e-16		
Matches 52	Conservative	0	Mismatches	0	Indels
				0	Gaps
				0	

TITLE	Direct Submission
JOURNAL	Submitted (01-SEP-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE	3 (bases 1 to 163059)
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., ... et al. (1999) The Human Genome Project. Nature 391: 315-321.

TITLE	Direct Submission
JOURNAL	Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
COMMENT	On Apr 8, 2000 this sequence version replaced gi:16630600.

```

http://ftp.genome.washington.edu/RM/RepeatMasker.html
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIMR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
Project Information
Center project name: 12031
Center clone name: 128_N 8
Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 139868 bases at least Q40
Consensus quality: 156241 bases at least Q30
Consensus quality: 162235 bases at least Q20
Insert size: 170000; agarose-fp
Insert size: 166559; sum-of-ctngs
Quality coverage: 3.9 in Q20 bases; agarose-fp
Quality coverage: 4.0 in Q20 bases; sum-of-ctngs
NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 1539: contig of 1539 bp in length
1540 1539: gap of 100 bp
1640 2821: contig of 1182 bp in length

```

```
* 2822 2921: gap of 100 bp
* 2922 4692: contig of 1771 bp in length
* 4693 4792: gap of 100 bp
* 4793 6136: contig of 1344 bp in length
* 6137 6236: gap of 100 bp
* 6237 8728: contig of 2492 bp in length
* 8729 8829: gap of 100 bp
* 8829 11075: contig of 2247 bp in length
* 11076 11175: gap of 100 bp
* 11176 13885: contig of 2710 bp in length
* 13886 13985: gap of 100 bp
* 13986 17062: contig of 3077 bp in length
* 17063 17162: gap of 100 bp
* 17163 21818: contig of 4656 bp in length
* 21819 25663: contig of 3745 bp in length
* 25664 25763: gap of 100 bp
* 25764 30441: contig of 4678 bp in length
* 30442 30541: gap of 100 bp
* 30542 33986: contig of 3345 bp in length
* 33987 33986: gap of 100 bp
* 33987 38474: contig of 4488 bp in length
* 38475 38574: gap of 100 bp
* 38575 43895: contig of 5321 bp in length
* 43896 43995: gap of 100 bp
* 43996 50044: contig of 6049 bp in length
* 50045 50144: gap of 100 bp
* 50145 56692: contig of 6548 bp in length
* 56693 56792: gap of 100 bp
* 56793 61378: contig of 4586 bp in length
* 61379 61478: gap of 100 bp
* 61479 69459: contig of 7981 bp in length
* 69460 69559: gap of 100 bp
* 69560 78975: contig of 9416 bp in length
* 78976 79075: gap of 100 bp
* 79076 88285: contig of 9210 bp in length
* 88286 88385: gap of 100 bp
* 88386 98036: contig of 9651 bp in length
* 98037 98136: gap of 100 bp
* 98137 109264: contig of 11128 bp in length
* 109265 109364: gap of 100 bp
* 109365 123064: contig of 13700 bp in length
* 123065 123164: gap of 100 bp
* 123165 143558: contig of 20394 bp in length
* 143559 143658: gap of 100 bp
* 143659 169059: contig of 25401 bp in length.
FEATURES
source
1. 169059
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="8"
/map="8"
/clone="RP11-128N8"
/clone_lib="RP11 Human Male BAC"
1. 1539
/note="assembly_fragment"
gap
1540. 1639
/estimated_length=100
misc_feature
1640. 2821
/note="assembly_fragment"
gap
2822. 2921
/estimated_length=100
misc_feature
2922. 4692
/note="assembly_fragment"
gap
4693. 4792
/estimated_length=100
misc_feature
4793. 6136
/note="assembly_fragment"
gap
6137. 6236
/estimated_length=100
misc_feature
6237. 8728
/note="assembly_fragment"
```

```
gap 8729. 8828
misc_feature /estimated_length=100
gap 8829. 11075
misc_feature /note="assembly_fragment"
gap 11076. 11175
misc_feature /estimated_length=100
gap 11176. 13885
misc_feature /note="assembly_fragment"
gap 13886. 13985
misc_feature /estimated_length=100
gap 13986. 17062
misc_feature /note="assembly_fragment"
gap 17063. 17162
misc_feature /estimated_length=100
gap 17163. 21818
misc_feature /note="assembly_fragment"
gap 21819. 25663
misc_feature /estimated_length=100
gap 25664. 25763
misc_feature /estimated_length=100
gap 25764. 30441
misc_feature /note="assembly_fragment"
gap 30442. 30541
misc_feature /estimated_length=100
gap 30542. 33986
misc_feature /note="assembly_fragment"
gap 33987. 38474
misc_feature /estimated_length=100
misc_feature /estimated_length=100
```

```
Query Match 1.7%; Score 52; DB 14; Length 169059;
Best Local Similarity 100.0%; Pred. No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3071 CAAGATTGGCCACGCACTCCAGCTGGGCAACAGAGACGCTGCTC 3122
DB 158812 CAAGATTGGCCACGCACTCCAGCTGGGCAACAGAGACGCTGCTC 158761
```

```
RESULT 146
AC021165/c 169893 bp DNA linear HTG 02-SEP-2000
LOCUS Homo sapiens chromosome 19 clone RP11-565J3, WORKING DRAFT
DEFINITION
ACCESSION AC021165
VERSION AC021165.4 GI:9965022
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 169893)
AUTHORS Waterston,R.H.
TITLES The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 169893)
AUTHORS Waterston,R.H.
TITLES Direct Submission
JOURNAL Submitted (14-JAN-2000) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Sep 2, 2000 this sequence version replaced gi:7637373.
```

```
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
```



```

Center project name: H_NH056503
----- Summary Statistics -----
Sequencing vector: M13, 761
Sequencing vector: plasmid, 241
Chemistry: Dye-primer RT; 761 of reads
Chemistry: Dye-terminator Big Dye, 241 of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 152885 bases at least Q40
Consensus quality: 159085 bases at least Q30
Consensus quality: 162352 bases at least Q20
Insert size: 185000; agarose-fp
Insert size: 168293; sum-of-contigs
Quality coverage: 3.07 in Q20 bases; agarose-fp
Quality coverage: 3.43 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of 'N', but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 1397: contig of 1397 bp in length
* 1398
* 1497: gap of unknown length
* 1498
* 3037: contig of 1540 bp in length
* 3038
* 3137: gap of unknown length
* 3138
* 6455: contig of 3318 bp in length
* 6456
* 6555: gap of unknown length
* 6556
* 9394: contig of 2839 bp in length
* 9395
* 9494: gap of unknown length
* 9495
* 12928: contig of 3434 bp in length
* 12929
* 13029
* 17141: contig of 4113 bp in length
* 17142
* 17241: gap of unknown length
* 17242
* 23019: contig of 5778 bp in length
* 23020
* 23119: gap of unknown length
* 23120
* 27396: contig of 4177 bp in length
* 27397
* 27396: gap of unknown length
* 27397
* 53901: contig of 26505 bp in length
* 53902
* 54001: gap of unknown length
* 54002
* 61627: contig of 7626 bp in length
* 61628
* 62127: gap of unknown length
* 61628
* 72148: contig of 10421 bp in length
* 72149
* 72248: gap of unknown length
* 72249
* 80663: contig of 8415 bp in length
* 80763
* 80763: gap of unknown length
* 80764
* 89249: contig of 8486 bp in length
* 89250
* 89349: gap of unknown length
* 89350
* 102831: contig of 13482 bp in length
* 102832
* 102931: gap of unknown length
* 102932
* 115578: contig of 12647 bp in length
* 115579
* 115678: gap of unknown length
* 115679
* 131546: contig of 15868 bp in length
* 131547
* 131647: gap of unknown length
* 131647
* 169893: contig of 38247 bp in length.
*
* Location/Qualifiers
*
* 1. 169893
* /organism="Homo sapiens"
* /mol_type="genomic DNA"
* /db_xref="taxon:9606"
* /chromosome="19"
* /clone="RP11-565J3"
*
* 1. 1397
* /note="assembly_name:Contig10"
*
* 1398. 1497
* /estimated_length=unknown
*
* 1498. 3037
* /note="assembly_name:Contig11"
*
* 3038. 3137
* /estimated_length=unknown
*
* 3138. 6455
* /note="assembly_name:Contig12"

```

Query Match	Best Local Similarity	1.7%; Score 52; DB 14; Length 169893;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0		
3071 CAAGATTGCGCACTGCACCTCCAGCTGGGCAACAGACCAAGACTCTGTCTC 3122		
DB 32826 CAAGTTGTGCGCACTGCACCTCCAGCTGGGCAACAGACCAAGACTCTGTCTC 32775		

LOCUS	AL135903	170232 bp	DNA	linear	PRI 18-MAY-2005
DEFINITION	Human DNA sequence from clone RPl1-30P6 on chromosome 6. Contains a Karyopherin alpha 5 (importin alpha 6) (KPNAS) pseudogene, a keratin 18 (KRT18) pseudogene, a tumor protein, translationally-controlled 1 (TP11) pseudogene, a pseudogene similar to dntp pyrophosphatase (DNT), the NT5E gene for 5'-nucleotidase ecto (CD73) and a Cpg island, complete sequence.				
ACCESSION	AL135903				
VERSION	AL135903.12	GI:12830388			
KEYWORDS	HTG; CD73; Cpg island; DNT; KPNAS; KRT18; NT5E; TP11.				
SOURCE	Human sapiens (human)				
ORGANISM	Human sapiens	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.			
REFERENCE	1 (bases 1 to 170232)	Johnson, C.			
AUTHORS	Direct Submission				
TITLE	Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk				
JOURNAL	On Feb 14, 2001 this sequence version replaced gi:11342762.				
COMMENT	The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information on the WORMPEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone configs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr6 RPl1-30P6 is from the library RPl1-11.1 constructed by the group of Pieter de Jong. For further details see http://www.ncbi.nlm.nih.gov/genbank/home.htm VECTOR: pBAC3.6				
	----- Genome Center				
	Center: Wellcome Trust Sanger Institute				
	Center code: SC				
	Web site: http://www.sanger.ac.uk				
	Contact: vegas@sanger.ac.uk				
	-----				
FEATURES	This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.				
SOURCE	Location/Qualifiers				
	1..170232				
	/organism="Homo sapiens"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:9606"				
	/chromosome="6"				
	/clone="RPl1-30P6"				
	/clone_lib="RPl1-11.1"				
	104				
	/note="Clone_right_end: RPl1-161C16"				
gene	join(520..610,682..1243,1488..2119,4690..4777,5077..5293)				
	/locus_tag="RPl1-30P6.1-001"				
	/pseudo				
CDS	join(520..610,682..1243,1488..2119,4690..4777,5077..5293)				
	/locus_tag="RPl1-30P6.1-001"				
	/note="match: proteins: AAL87378 IMA1 SCPO IMA2 ARATH IMA3 HUMAN IMA LYCES KIM8 HUMAN K2C8 MOUSE O49601 O82783 O94374 Q90YH3 Q94F55 Q9D637 Q9PUB5 Q9UIH6"				
	/pseudo				
	/codon_start=1				
gene	2233..3451				
	/locus_tag="RPl1-30P6.2-001"				
	/pseudo				
CDS	2233..3451				
	/locus_tag="RPl1-30P6.2-001"				

gene	/note="match: proteins: AAB04253 AAB00982 KICR HUMAN K1CR_MOUSE K1M8 HUMAN K2C8_MOUSE OS7611 077727 Q07427 Q90M74 Q90YH3 Q96GD2 Q9PUB5 Q9PWB8 Q9QWL7"				
	/pseudo				
	/codon_start=1				
mRNA	complement(join(101663..104001,104457..104630))				
	/locus_tag="RPl1-30P6.6-001"				
	/pseudo				
gene	complement(join(101663..104001,104457..104630))				
	/locus_tag="RPl1-30P6.6-001"				
	/note="match: cDNAs: AK024998"				
	complement(141101..141569)				
gene	/locus_tag="RPl1-30P6.4-001"				
	/pseudo				
CDS	complement(141101..141569)				
	/locus_tag="RPl1-30P6.4-001"				
	/note="match: proteins: AAB96413 AAD50334 AAL77806 CAB54626 DUT ADEG1 DUT ADEB8 DUT FOWPV DUT HUMAN DUT OR DUT RAT DUT SCPO DUT VACC1 DUT VARV DUT YEAST O01934 P87530 Q8V546 Q8VCG1 Q91BK5 Q96Q81 Q99HX3 Q9CQ43 Q9CU930 Q9URF1 Q9UJ44 Q9Q8S9 Q9Q958 Q9STG6 Q9UI74"				
	/pseudo				
	/codon_start=1				
gene	complement(141897..142410)				
	/locus_tag="RPl1-30P6.3-001"				
	/pseudo				
CDS	complement(141897..142410)				
	/locus_tag="RPl1-30P6.3-001"				
	/note="match: proteins: AAI22436 AAL68965 PKG2 HUMAN Q9NWX1 TC1P BRAB TC1P CHICK TC1P DRMB TC1P HUMAN TC1P LABRO TC1P MEDSA TC1P MOUSE TC1P_PEA TC1P_RABIT TC1P_SCHUA TC1P_SCPO TC1P_YEAST"				
	/pseudo				
	/codon_start=1				
gene	join(164535..164922,AL589666.5:11372..11594,AL589666.5:15549..15737,AL589666.5:29547..29744,AL589666.5:31647..31801,AL589666.5:33806..33911,AL589666.5:34820..34969,AL589666.5:36289..36489,AL589666.5:38153..40090)				
	/gene="NT5B"				
	/locus_tag="RPl1-321N4.1-001"				
	join(164535..164922,AL589666.5:11372..11594,AL589666.5:15549..15911)				
gene	/gene="NT5B"				
	/locus_tag="RPl1-321N4.1-002"				
	join(164535..164922,AL589666.5:11372..11594,AL589666.5:15549..15911)				
	/gene="NT5B"				
	/locus_tag="RPl1-321N4.1-001"				
	/product="5'-nucleotidase, ecto (CD73)"				
	/note="match: ESTs: AI202304 A1420969 AW006723 AW025554 AW271686 BE677446 BF476532				
mRNA	match: cDNAs: BC015940 BC0015940.1"				
	join(164535..164922,AL589666.5:11372..11594,AL589666.5:15549..15737,AL589666.5:29547..29744,AL589666.5:31647..31801,AL589666.5:33806..33911,AL589666.5:34820..34969,AL589666.5:36289..36489,AL589666.5:38153..40090)				
	/gene="NT5B"				
	/locus_tag="RPl1-321N4.1-001"				
	/product="5'-nucleotidase, ecto (CD73)"				
	/note="match: ESTs: AA375089 AL553737 AL702918 BG646149 BE547456 BB891902 BG294369 B1756281 BM006258 BM762339 BM847311 BM926263 BM937597				
CDS	match: cDNAs: BT5NUC D05NUCLT H55NUSE MME5NA RNSRHP S64302"				
	join(164584..164922,AL589666.5:11372..11594,AL589666.5:15549..15781)				
	/gene="NT5B"				
	/locus_tag="RPl1-321N4.1-001"				
	/product="5'-nucleotidase, ecto (CD73)"				
	/note="match: ESTs: AA375089 AL553737 AL702918 BG646149 BE547456 BB891902 BG294369 B1756281 BM006258 BM762339 BM847311 BM926263 BM937597				
	match: cDNAs: BT5NUC D05NUCLT H55NUSE MME5NA RNSRHP S64302"				
	join(164584..164922,AL589666.5:11372..11594,AL589666.5:15549..15781)				
	/gene="NT5B"				
	/locus_tag="RPl1-321N4.1-001"				
	/standard_name="OTTHUMP0000016809"				
	/note="match: proteins: Q95B60"				
	/codon_start=1				
	/product="5'-nucleotidase, ecto (CD73)"				

/note="match: proteins: AAH04253 AAH20982 KICR HUMAN KICR MOUSE KIM8 HUMAN K2C8 MOUSE O57611 O77727 O07427 Q90W74 Q90YH3 Q96GD2 Q9PUB5 Q9PWB8 Q9QML7"

/pseudo

/codon\_start=1

/complement(join(101663..104001,104457..104630))

/locus\_tag="RPl1-30P6.6-001"

/complement(join(101663..104001,104457..104630))

/locus\_tag="RPl1-30P6.6-001"

/note="match: cDNAs: AK024998"

/complement(join(141101..141569))

/locus\_tag="RPl1-30P6.4-001"

/pseudo

/complement(join(141101..141569))

/locus\_tag="RPl1-30P6.4-001"

/note="match: proteins: AAB86413 AAD50334 AAL77806

CAB54626 DUT\_ADEG1 DUT\_ADEG8 DUT\_FOWPV DUT\_HUMAN DUT\_ORFN2

DUT\_RAT DUT\_SCHPO DUT\_VACC DUT\_VARY DUT\_YEAST O01934

P87630 Q8V46 Q8VCG1 Q91BK5 Q96Q81 Q99HX3 Q9CC43 Q9CU90

Q9JF1 Q9J744 Q9Q8S9 Q9Q958 Q9S7G6 Q9UI74"

/pseudo

/codon\_start=1

/complement(join(141897..142410))

/locus\_tag="RPl1-30P6.3-001"

/pseudo

/complement(join(141897..142410))

/locus\_tag="RPl1-30P6.3-001"

/note="match: proteins: AAL22436 AAL68965 FKZ2 HUMAN

Q9NMV1 TCTP\_BRABR TCTP\_CHICK TCTP\_DROME TCTP\_HUMAN

TCTP\_LABRO TCTP\_MEDSA TCTP\_MOUSE TCTP\_PEA TCTP\_RABIT

TCTP\_SCHUA TCTP\_SCHPO TCTP\_YEAST"

/pseudo

/codon\_start=1

/join(164535..164922,AL589666.5:11372..11594,

AL589666.5:15549..15737,AL589666.5:29547..29744,

AL589666.5:31647..31801,AL589666.5:33806..33911,

AL589666.5:34820..34969,AL589666.5:36289..36489,

AL589666.5:38153..40090)

/gene="NT5E"

/locus\_tag="RPl1-321N4.1-001"

/join(164535..164922,AL589666.5:11372..11594,

AL589666.5:15549..15737,AL589666.5:29547..29744,

AL589666.5:31647..31801,AL589666.5:33806..33911,

AL589666.5:34820..34969,AL589666.5:36289..36489,

AL589666.5:38153..40090)

/gene="NT5E"

/locus\_tag="RPl1-321N4.1-001"

/product="5'-nucleotidase, ecto (CD73)"

/note="match: ESTs: A1202304 A1420969 AM006723 AM025554

AW271686 BE677446 BF476532

match: cDNAs: BC015940 BC015940.1"

/join(164535..164922,AL589666.5:11372..11594,

AL589666.5:15549..15737,AL589666.5:29547..29744,

AL589666.5:31647..31801,AL589666.5:33806..33911,

AL589666.5:34820..34969,AL589666.5:36289..36489,

AL589666.5:38153..40090)

/gene="NT5E"

/locus\_tag="RPl1-321N4.1-001"

/product="5'-nucleotidase, ecto (CD73)"

/note="match: ESTs: AA375088 AL553737 AL702918 BB646149

BE547456 BR891902 BG924369 B1756281 BM006258 BM762339

BM847311 BM926263 BM937597

match: cDNAs: BT5NOC D05NUCT H55NUASE MME5NA RN5RPH

S64302"

/join(164584..164922,AL589666.5:11372..11594,

AL589666.5:15549..15781)

/gene="NT5E"

/locus\_tag="RPl1-321N4.1-001"

/standard\_name="OTTHUMP0000016809"

/note="match: proteins: Q96B60"

/codon\_start=1

/product="5'-nucleotidase, ecto (CD73)"

```

/protein_id="CAH72336.1"
/db_xref="GI:55661713"
/db_xref="GOA:Q96860"
/db_xref="InterPro:IPR004843"
/db_xref="InterPro:IPR006146"
/db_xref="InterPro:IPR006179"
/db_xref="UniProt/TrEMBL:Q96860"
/translation="MCPRAAPABTLLALGAVLMPAPAGAWETLTHNDVSRLEQT
SDSSKCVNASCWGVARLPTKVOQIRAPENYVLLDAGDOYQTIPTFYKGAFA
HFMALRYDAMALGNHEDFNGVGLIPLKAKPPIISANTKAGPLASQISGLYLP
YVLPVDEVGVIIGVTSKETPELSNPGTILAFEBEITLALQEVKLTILNNKIAL
GSGFEMDGLIAQKRGVGVVGGSNFTLVGNPSPKSPKAPKPTVTSDDGRKP
VVOAARFGVLYGLIKTIBEDRGNAVISHENPILNLSIPEDSIPADINKKIKIDNG
STOBEKATVYIDSSQSCRFECNKNILCDAMINNNLRHIDEMFAMHVSIGSTSEFL
GSRGPIDEKNNGITITMENLAALPFGGTFDLVQLKSGTDLKAFHSHVHVGSTSEFL
QVGGIHVVYLDLSRKQDRVYKLDVLCIKGRVPSYDPLNNDEVYKYLINPLANGDGF
QMKIDELRHDSGDQDINVSITYISKMKVTPAVVGRKIFSTGSHSGSFSLIFLSLW
AIVFLVYC"
/misc_feature
170232
/note="Clone_right_end: RP11-30P6"

Query Match      1.7%; Score 52; DB 8; Length 170232;
Best Local Similarity 100.0%; Pred. No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTTGCTCTC 3122
      |||||
Db 128510 CAAGATTGGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTTGCTCTC 128459

RESULT 148
AC073068 171058 bp DNA linear PRI 08-OCT-2003
DEFINITION Homo sapiens BAC clone RP11-159118 from 7, complete sequence.
ACCESSION AC073068
VERSION AC073068.8 GI:14010911
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 171058)
Suleton,J.E. and Wilson,R.
Toward a complete human genome sequence
Genome Res. 8 (11), 1097-1108 (1998)
PUBMED 9847074

```

```

REFERENCE
AUTHORS 2 (bases 1 to 171058)
          Belter,E., Maupin,R. and Podhrasky,A.
TITLE The sequence of Homo sapiens BAC clone RP11-159118
JOURNAL Unpublished (2001)
REFERENCE
AUTHORS 3 (bases 1 to 171058)
          Wacziarg,R.H.
TITLE Direct Submission
JOURNAL Submitted (08-JUN-2000) Genome Sequencing Center, Washington
          University School of Medicine, 4444 Forest Park Parkway, St. Louis,
          MO 63108, USA
REFERENCE
AUTHORS 4 (bases 1 to 171058)
          Wacziarg,R.
TITLE Direct Submission
JOURNAL Submitted (09-MAY-2001) Department of Genetics, Washington
          University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE
AUTHORS 5 (bases 1 to 171058)
          Wacziarg,R.
TITLE Direct Submission
JOURNAL Submitted (29-APR-2003) Department of Genetics, Washington
          University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
          6 (bases 1 to 171058)
          Wilson,R.
REFERENCE
AUTHORS Direct Submission
          Submitted (08-OCT-2003) Department of Genetics, Washington
          University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
          On May 9, 2001 this sequence version replaced gi:10518416.
COMMENT ----- Genome Center
          Center: Washington University Genome Sequencing Center
          Center code: WUGSC
          Web site: http://genome.wustl.edu
          Contact: saplen@wustl.wustl.edu
          ----- Summary Statistics
          -----
          Center project name: H_NH0159118
          -----

NOTICE: This sequence may not represent the entire insert of this
clone. It may be shorter because we only sequence overlapping
clone sections once, or longer because we provide a small overlap
between neighboring data submissions.

This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
The sequence of this clone was established as part of a mapping and
sequencing collaboration between the NHGRI Chromosome 7 Mapping
Project (Eric D. Green, Director), John D. McPherson in the
Department of Genetics (Washington University), and the Washington
University Genome Sequencing Center. For additional information
about the map position of this sequence, see
http://www.nhgri.nih.gov/DIR/GRB/CHR7, send
mailto:egreen@nhgri.nih.gov, or see http://genome.wustl.edu

SOURCE INFORMATION:
The RP11-11 human BAC library was made from the blood of one male
donor, as described by Osoegawa,K., Woon,P.Y., Zhao,B., Frenken,E.,
Tateo,M., Catanesi,J.V., and de Jong,P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.chori.org
VECTOR: PBAC3.6

NEIGHBORING SEQUENCE INFORMATION:
Actual start of this clone is at base position 1 of RP11-159118
actual end is at base position 171058 of RP11-159118.
Location/Qualifiers
1..171058

```

```
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="7"
/map="7"
/clone="RP11-159118"
/clone_1fb="RPCT-11"
1. 835
/rpt_family="L1"
repeat_region
/rpt_family="L1"
839. 1818
/rpt_family="L1"
repeat_region
/rpt_family="ERV1"
2300. 2408
/rpt_family="ERV1"
2997. 3302
/rpt_family="Alu"
repeat_region
/rpt_family="AT-rich"
3483. 3512
/rpt_family="AT-rich"
5688. 5842
/rpt_family="L1"
6268. 6676
/rpt_family="L1"
6693. 7000
/rpt_family="Alu"
7003. 7114
/rpt_family="L1"
7151. 7315
/rpt_family="L1"
7316. 7618
/rpt_family="Alu"
7619. 7838
/rpt_family="L1"
8030. 8166
/rpt_family="GA-rich"
8172. 8986
/rpt_family="L1"
8986. 9561
/rpt_family="L1"
9574. 9723
/rpt_family="GA-rich"
9809. 10484
/rpt_family="L1"
10524. 10552
/rpt_family="AT-rich"
10745. 10968
/rpt_family="L1"
10975. 11399
/rpt_family="L1"
11705. 11773
/rpt_family="L1"
11783. 11985
/rpt_family="Alu"
11986. 12013
/rpt_family="L1"
12014. 12037
/rpt_family="(TG)n"
12038. 13301
/rpt_family="L1"
13302. 13594
/rpt_family="Alu"
13595. 14332
/rpt_family="L1"
14331. 14700
/rpt_family="L1"
16105. 16181
/rpt_family="MER1_type"
17251. 17343
/rpt_family="ERV1"
17360. 17741
/rpt_family="ERV1"
18632. 18662
/rpt_family="(TTCC)n"
18657. 18700
/rpt_family="CT-rich"

repeat_region 19609. 19720
/rpt_family="MIR"
repeat_region 20991. 21033
/rpt_family="(TG)n"
repeat_region 21046. 21083
/rpt_family="(TTTG)n"
repeat_region 21088. 21271
/rpt_family="Alu"
repeat_region 21754. 21916
/rpt_family="MIR"
repeat_region 22245. 22279
/rpt_family="CRI"
repeat_region 22735. 23580
/rpt_family="L1"
repeat_region 24125. 25030
/rpt_family="L1"
repeat_region 25031. 25336
/rpt_family="Alu"
repeat_region 25337. 25546
/rpt_family="L1"
repeat_region 25550. 25632
/rpt_family="Mariner"
repeat_region 25663. 25997
/rpt_family="ERV1"
repeat_region 26166. 26337
/rpt_family="MALR"
repeat_region 26338. 26431
/rpt_family="ERV1"
repeat_region 26441. 27629
/rpt_family="L1"
repeat_region 27630. 27683
/rpt_family="L1"
repeat_region 27684. 27970
/rpt_family="ERV1"
repeat_region 27993. 28276
/rpt_family="L1"
repeat_region 28315. 28609
/rpt_family="L1"
repeat_region 29266. 29353

Query Match 1.7%; Score 52; DB 8; Length 171058;
Best Local Similarity 100.0%; Pred. No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCGACCTGCACTCGAGCTTGCGCAACAGAGCAAGACTGTCTC 3122
DB 7545 CAAGATTGTGCGACCTGCACTCGAGCTTGCGCAACAGAGCAAGACTGTCTC 7596

RESULT 149
AC010616 171849 bp DNA linear PRI 01-MAY-2001
LOCUS Homo sapiens chromosome 19 clone CTD-2575K13, complete sequence.
DEFINITION AC010616
ACCESSION AC010616.5 GI:13876489
VERSION HTG.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 171849)
DOE Joint Genome Initiative and Stanford Human Genome Center.
AUTHORS Direct Submission
TITLES Unpublished
JOURNAL
REFERENCE 2 (bases 1 to 171849)
DOE Joint Genome Initiative.
AUTHORS Direct Submission
TITLE Submitted (16-SEP-1999) Production Sequencing Facility, DOE Joint
JOURNAL Genome Initiative, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE 3 (bases 1 to 171849)
DOE Joint Genome Initiative and Stanford Human Genome Center.
AUTHORS Direct Submission
TITLE
```

JOURNAL	Submitted (28-APR-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA			
REFERENCE	4 (bases 1 to 171849)			
AUTHORS	DOE Joint Genome Institute and Stanford Human Genome Center.			
TITLE	Direct Submission			
JOURNAL	Submitted (01-MAY-2001) DOE Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA			
COMMENT	On Apr 28, 2001 this sequence version replaced gi:7711477. Draft Sequence Produced by DOE Joint Genome Institute www.jgi.doe.gov Finishing Completed at Stanford Human Genome Center www-shgc.stanford.edu Quality: Phrap Quality >=40 99.8% of Sequence; Estimated Total Number of Errors is 1. STS Content: WT-15934 G21486 SHGC-64254 G38544. Location/Qualifiers 1..171849 /organism="Homo sapiens" /mol_type="genomic DNA" /db_xref="taxon:9606" /chromosome="19" /clone="CID-2575K13"			
FEATURES				
SOURCE				
ORIGIN				
Query Match	1.7%	Score 52;	DB 8;	Length 171849;
Best Local Similarity	100.0%;	Pred. No. 5,7e-16;		
Matches	52;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
Oy	3071	CAAGATTGTGGCCTGCACCTCCAGCTGGGCAACAGACGACCTCTCTTC	3122	
Db	129747	CAAGATTGTGGCCTGCACCTCCAGCTGGGCAACAGACGACCTCTCTTC	129798	
RESULT 150				
LOCUS	AL590609/c			
DEFINITION	AL590609 172307 bp DNA linear PRI 18-MAY-2005			
	Human DNA sequence from clone RP11-4M23 on chromosome 1 Contains the 5' end of the FUC1 gene for alpha-L-1 fucosidase (tissue) the CR2 gene for cannabinoid receptor 2 (macrophage), a BTB (POZ) domain-containing 6 (BTBD6) pseudogene, a novel gene, an H3 histone family 3A (H3F3A) pseudogene, the PNR2 gene for prolins-rich nuclear receptor coregulator 2, the FUS1P1 gene for FUS interacting protein (serine-arginine rich) 1, a ribosomal protein L36 (RPL36) pseudogene and seven CpG islands, complete sequence.			
ACCESSION	AL590609			
KEYWORDS	AL590609.15	GI:395773520		
VERSIONS	HTG; BTBD6; cannabinoid receptor; CR2; CpG island; FUC1A; FUS; FUS1P1; H3F3A; histone; PNR2; POZ; prolins-rich; RPL36.			
SOURCE	Homo sapiens (human)			
ORGANISM	Homo sapiens			
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.			
REFERENCE	1 (bases 1 to 172307)			
AUTHORS	Wallis, J.			
TITLE	Direct Submission			
JOURNAL	Submitted (13-MAY-2005) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: vegas@sanger.ac.uk			
COMMENT	Clone requests: clonerequest@sanger.ac.uk On Dec 8, 2003 this sequence version replaced gi:15795484. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em; EMBL; Sw; SWISSPROT; Tr; TREMBL; Wp; WORMEP; Information on the WORMEP database can be found at http://www.sanger.ac.uk/Projects/C_elegans/wormep This sequence was generated from part of bacterial clone configs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chr1 RP11-4M23 is from the library RP11-11.1 constructed by the group of Pieter de Jong. For further details see			

```

http://www.chemj.9
VECTOR: pBACe3.6
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: vega@sanger.ac.uk
-----
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one subclone; and the assembly was confirmed by restriction digest,
except on the rare occasion of the clone being a YAC.
Location/Qualifiers
1. 172307
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-4M23"
/clone_1hb="RP11-11.1"
1
/note="Clone left end: RP11-4M23"
complement(join(3548..8684,46268..46384))
/gene="CNR2"
/locus_tag="RP11-4M23.1-001"
complement(join(3548..8684,46268..46384))
/gene="CNR2"
/locus_tag="RP11-4M23.1-001"
/product="cannabinoid receptor 2 (macrophage)"
/note="match: ESTs: Em:AA485015.1 Em:AA509480.1
Em:A1218261.1 Em:A142494.1 Em:A1475798.1 Em:A1864752.1
Em:AW503816.1 Em:BM151899.1
match: cDNAs: Em:X74328.1"
complement(3548)
/gene="CNR2"
/locus_tag="RP11-4M23.1-001"
complement(3566..3571)
/gene="CNR2"
/locus_tag="RP11-4M23.1-001"
complement(6097)
/gene="CNR2"
/locus_tag="RP11-4M23.1-001"
complement(6116..6121)
/gene="CNR2"
/locus_tag="RP11-4M23.1-001"
complement(7557..8639)
/gene="CNR2"
/locus_tag="RP11-4M23.1-001"
/standard_name="OTTHUMP000000015772"
/note="match: proteins: Sw:PF34972 Sw:PF47936 Sw:Q9QZ9N
Tr:BAIC29520 Tr:BAIC29894"
/codon_start=1
/product="cannabinoid receptor 2 (macrophage)"
/protein_id="CA14799.1"
/db_xref="GI:55961036"
/db_xref="InterPro:IPR000276"
/db_xref="InterPro:IPR001551"
/db_xref="InterPro:IPR002230"
/translation="MECWTEIANGSKDGLDSNPMKDYMLSGPQRTAAVALVLTLLG
LISLAEVALVILISGHSRKRPSYFISGLAGADFLASVFAESFVNPHVFGVDS
KAYFLIKGVITMTFTASVGSLLITADRYLCIRPSPYVALIRGALVLTGLMTVL
SAVSYPLMGWTCPCPCESEPLIPNDYLSWLEPIAFISGIITGVHVMKAO
HVASLSGSDRQVPGMARERLDVRIATLGLVLAVALIICFPIVALMAHSLATLTSQ
VKAFAFCSMILCLINSVNVPIYALRGEIRSSAHCLAMKCKVRGLGSAKEAER
SVKTEITADKTIIPWPSRDLDSDC"
complement(34493..35759)
/locus_tag="RP11-4M23.2-001"
/pseudo
complement(34493..35759)
/locus_tag="RP11-4M23.2-001"

```

```

/note="match: PROTEIN: Sw:PS8545 Sw:Q96KX9 Sw:Q9Y2Y9
Tr:AAH62968 Tr:Q8CHC1 Tr:Q9NST6"
/pseudo
/codon_start=1
/product="BTB (POZ) domain-containing 6 (BTBD6)
pseudogene"
/complement(join(40133..40203,40769..40907))
/locus_tag="RP11-4M23.3-001"
/complement(join(40133..40203,40769..40907))
/locus_tag="RP11-4M23.3-001"
/product="proline-rich nuclear receptor coactivator 2"
/note="match: ESTs: Em:CD696420.1"
/73316
/note="Clone right end: RP11-45G17"
/complement(82037..82405)
/locus_tag="RP11-4M23.4-001"
/pseudo
/complement(82037..82405)
/locus_tag="RP11-4M23.4-001"
/locus_tag="RP11-4M23.4-001"
/note="match: Proteins: Sw:Q92068 Tr:AAH12687 Tr:AAH17558
Tr:BAH22464 Tr:BAC29895 Tr:BAC40130 Tr:Q7ZV67 Tr:Q8WSF1
Tr:Q9U281"
/pseudo
/codon_start=1
/product="H3 histone, family 3A (H3F3A) pseudogene"
/join(92131..92278,93951..94156,94462..94487)
/gene="PnRC2"
/locus_tag="RP11-4M23.5-003"
/join(92131..92278,93951..94156,94462..94487)
/gene="PnRC2"
/locus_tag="RP11-4M23.5-003"
/product="proline-rich nuclear receptor coactivator 2"
/note="match: ESTs: Em:AV686572.1 Em:CD689388.1"
/join(92834..93020,93951..94156,94462..96484)
/gene="PnRC2"
/locus_tag="RP11-4M23.5-001"
/join(92834..93020,93951..94156,94462..96484)
/gene="PnRC2"
/locus_tag="RP11-4M23.5-001"
/product="proline-rich nuclear receptor coactivator 2"
/note="match: ESTs: Em:A1221857.1 Em:A1240994.1
Em:A1684547.1 Em:A1823428.1 Em:A1829468.1 Em:A0280480.1
Em:A1840427.1 Em:AM274263.1 Em:BE047281.1 Em:BM922437.1
Em:BU159988.1 Em:BU165438.1
match: cDNAs: Em:AP151042.1 Em:AF374386.1 Em:AK000319.1
Em:AK024774.1 Em:BC001959.1"
/join(92850..93020,93955..94156,94462..94902)
/gene="PnRC2"
/locus_tag="RP11-4M23.5-002"
/join(92850..93020,93955..94156,94462..94902)
/gene="PnRC2"
/locus_tag="RP11-4M23.5-002"
/product="proline-rich nuclear receptor coactivator 2"
/note="match: ESTs: Em:BG15201.1 Em:BI596502.1"
94480..94899
/gene="PnRC2"
/locus_tag="RP11-4M23.5-001"
/standard_name="ORTHOMP0000015770"
/note="match: proteins: Tr:Q9CRJ3 Tr:Q9CKC6 Tr:Q9NPJ4"
/codon_start=1
/product="proline-rich nuclear receptor coactivator 2"
/protein_id="CA114802.1"
/db_xref="GI:55961037"
/db_xref="GOA:Q9NPJ4"
/db_xref="UniProt:Q9NPJ4"
/translation="MGGGRINYPAPQSNVSNKQOOLNRQTKESQNSOKTIYHKKKE
RHHGQNSAAQAQMGQNGKNNPNNQSNWSLSLGGPRLFRSQNQNQYAAKSEEPPE
SPSVLPKPSHPHVPVPSFNPSDEKIMTFQKTLTKIQV"
95681..95686
/gene="PnRC2"
/locus_tag="RP11-4M23.5-001"
95703
/gene="PnRC2"

```

polyA\_signal  
/locus\_tag="RP11-4M23.5-001"  
96441..96446  
/gene="PNRC2"  
/locus\_tag="RP11-4M23.5-001"  
96445..96450  
/gene="PNRC2"  
/locus\_tag="RP11-4M23.5-001"  
96484  
/gene="PNRC2"  
/locus\_tag="RP11-4M23.5-001"  
complement(join(97826..100745,104595..104648,  
104872..105034,107996..108099,111735..111839,  
113213..113378))  
/gene="FUSIP1"  
/locus\_tag="RP11-4M23.6-002"  
complement(join(97826..100745,104595..104648,  
104872..105034,107996..108099,111735..111839,  
113213..113378))

polyA\_signal  
/locus\_tag="RP11-4M23.5-001"  
96441..96446  
/gene="PNRC2"  
/locus\_tag="RP11-4M23.5-001"  
96445..96450  
/gene="PNRC2"  
/locus\_tag="RP11-4M23.5-001"  
96484  
/gene="PNRC2"  
/locus\_tag="RP11-4M23.5-001"  
complement(join(97826..100745,104595..104648,  
104872..105034,107996..108099,111735..111839,  
113213..113378))  
/gene="FUSIP1"  
/locus\_tag="RP11-4M23.6-002"  
complement(join(97826..100745,104595..104648,  
104872..105034,107996..108099,111735..111839,  
113213..113378))

polyA\_site  
96484  
/gene="PNRC2"  
/locus\_tag="RP11-4M23.5-001"  
complement(join(97826..100745,104595..104648,  
104872..105034,107996..108099,111735..111839,  
113213..113378))  
/gene="FUSIP1"  
/locus\_tag="RP11-4M23.6-002"  
complement(join(97826..100745,104595..104648,  
104872..105034,107996..108099,111735..111839,  
113213..113378))

Query Match 1.7%; Score 52; DB 8; Length 172307;  
Best Local Similarity 100.0%; Pred. No. 5,7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCATTCGATTCGACCTCGGGCAACAGACGAAGACTCTGTCTC 3122  
|||  
Db 149550 CAAGATTGTGCATTCGATTCGACCTCGGGCAACAGACGAAGACTCTGTCTC 149499

RESULT 151

AC025589  
LOCUS  
DEFINITION Homo sapiens chromosome 3 clone RP11-175f9, complete sequence.  
ACCESSION AC025589  
VERSION AC025589.20 GI:13899418  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
1 (bases 1 to 172759)  
Abola,A.P., Bruno,D., Conn,L., Dela Rosa,M., Faulkner,D.,  
Federpiel,N., Glukhov,S., Hansen,N., Herman,Z.S., Hyman,R.,  
Mao,J., Komp,C., Kotler,S., Lam,B., Marathe,R., Miranda,M.,  
Morehouse,A.J., Nguyen,M., Oefner,P., Palm,C.J., Ramirez,D.,  
Southwick,A.M., Webb,C., Wilhelmy,J., Yu,S. and Davis,R.W.  
Unpublished  
2 (bases 1 to 172759)  
Bruno,D., Conn,L., Dela Rosa,M., Faulkner,D., Federpiel,N.,  
Glukhov,S., Hansen,N., Hyman,R., Mao,J., Marathe,R.,  
Morehouse,A.J., Oefner,P., Palm,C.J., Ramirez,D.,  
Yu,S. and Davis,R.W.  
Direct Submission  
Submitted (12-MAR-2000) DNA Sequencing and Technology Center,  
Stanford University, 855 California Avenue, Palo Alto, CA 94304,  
USA  
3 (bases 1 to 172759)  
Abola,A.P., Bruno,D., Conn,L., Dela Rosa,M., Faulkner,D.,  
Federpiel,N., Glukhov,S., Hansen,N., Herman,Z.S., Hyman,R.,  
Komp,C., Kotler,S., Lam,B., Mao,J., Marathe,R., Miranda,M.,  
Morehouse,A.J., Nguyen,M., Oefner,P., Palm,C.J., Ramirez,D.,  
Southwick,A.M., Webb,C., Wilhelmy,J., Yu,S. and Davis,R.W.  
Direct Submission  
Submitted (01-MAY-2001) Genome Technology Center, Stanford  
University, 855 California Avenue, Palo Alto, CA 94304, USA  
On May 1, 2001 this sequence version replaced gi:113446262.  
----- Genome Center  
Center: Stanford DNA Sequencing and Technology Development  
Center  
Center code: SDSSTD  
Web site: http://sequence-www.stanford.edu/group/human/  
Contact: hum-info@sequence.stanford.edu  
----- Project Information



Center project name: 860  
Center clone name: RP11-175F9

## ----- Summary

This sequence was finished as follows unless otherwise noted:  
all regions were double stranded or sequenced with an alternate  
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such  
as compressions and repeats; all regions were covered by sequence  
from more than one subclone; and the assembly was confirmed by  
restriction digest.

## FEATURES

source

1..172759

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="3"

/clone="RP11-175F9"

/clone\_11b="RPC1 human BMC library 11"

## ORIGIN

Query Match 1.7%; Score 52; DB 8; Length 172759;  
Best Local Similarity 100.0%; Pred. No. 5.7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2899 GATCACCCTGAGCGCCAGAGTTCGAGACGACCTGGCCACATAGCGAACC 2350

DB 14184 GATCACCCTGAGCGCCAGAGTTCGAGACGACCTGGCCACATAGCGAACC 14235

## RESULT 152

AC010160/c

LOCUS

DEFINITION

SEQUENCE

AC010160

AC010160.9

VERSION

KEYWORDS

HTG; HTGS PHASE1; HTGS\_DRAFT; HTGS\_FULLTOP; HTGS\_ACTIVEFIN;

HTGS\_CANCELLED.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homnidae; Homo.

1 (bases 1 to 172915)

Smith, D.R.

Genome Therapeutics Corporation Sequencing Center: Human Genome

Sequence Data

Unpublished

2 (bases 1 to 172915)

Smith, D.R.

Direct Submission

Submitted (14-SEP-1999) Genome Therapeutics Corporation, 100 Beaver

Street, Waltham, MA 02453, USA

On May 6, 2001 this sequence version replaced gi:9887635.

COMMENT

Genome Center

Center: Genome Therapeutics Corporation

Center code: GTC

Web site: http://www.genomecorp.com/

Contact: gtc-seqcenter@genomecorp.com

Project Information

Center project name: hg015

Summary Statistics

Sequencing vector: N/A

Chemistry: Dye-terminator Big Dye, 100% of reads

Assembly program: Phrap; version 990315

Consensus quality: 166031 bases at least Q40

Consensus quality: 167418 bases at least Q30

Consensus quality: 168334 bases at least Q20

Insert size: 172565; sum-of-contigs

Quality coverage: 7.4x in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently

\* consists of 5 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1479: contig of 1479 bp in length  
\* 1480 1579: gap of unknown length  
\* 1580 19522: contig of 17943 bp in length  
\* 19523 19622: gap of unknown length  
\* 19623 43164: contig of 23542 bp in length  
\* 43165 43264: gap of unknown length  
\* 43265 88435: contig of 45171 bp in length  
\* 88436 88536: gap of unknown length  
\* 88536 172915: contig of 84380 bp in length.

## FEATURES

source

1..172915

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="10"

/clone="RP11-396M20"

/clone\_11b="RPC1-11"

1..1479

/note="assembly\_name:Contig3"

1480..1579

/estimated\_length=unknown

1580..19522

/note="assembly\_name:Contig5"

19523..19622

/estimated\_length=unknown

19623..43164

/note="assembly\_name:Contig6"

43165..43264

/estimated\_length=unknown

43265..88435

/note="assembly\_name:Contig7"

88436..88535

/estimated\_length=unknown

88536..172915

/note="assembly\_name:Contig8"

## ORIGIN

Query Match 1.7%; Score 52; DB 14; Length 172915;  
Best Local Similarity 100.0%; Pred. No. 5.7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCACCTGCTCCAGCTGGGCAACAGCAAGACTCTGCTC 3122

DB 153187 CAAGATTGTGCACCTGCTCCAGCTGGGCAACAGCAAGACTCTGCTC 153136

## RESULT 153

AC020908

LOCUS

DEFINITION

AC020908

AC020908.7

VERSION

KEYWORDS

HTG.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homnidae; Homo.

1 (bases 1 to 174034)

DOE Joint Genome Institute and Stanford Human Genome Center.

Direct Submission

Unpublished

2 (bases 1 to 174034)

DOE Joint Genome Institute.

Direct Submission

Submitted (12-JAN-2000) Production Sequencing Facility, DOE Joint

REFERENCE 3 (bases 1 to 174034)  
 DOB Joint Genome Institute and Stanford Human Genome Center.  
 TITLE Direct Submission  
 JOURNAL Submitted (01-SEP-2000) DOB Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 REFERENCE 4 (bases 1 to 174034)  
 DOB Joint Genome Institute and Stanford Human Genome Center.  
 TITLE Direct Submission  
 JOURNAL Submitted (03-OCT-2001) DOB Joint Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
 COMMENT On Oct 3, 2001 this sequence version replaced gi:9958017.  
 Draft Sequence Produced by DOB Joint Genome Institute  
 www.jgi.doe.gov  
 Finishing Completed at Stanford Human Genome Center  
 www.shgc.stanford.edu  
 Quality: Phrap Quality >=40 99.8% of Sequence;  
 Estimated Total Number of Errors is 0.6.  
 Location/Qualifiers  
 1. 174034  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="19"  
 /clone="CTD-2528A14"

ORIGIN  
 Query Match 1.7%; Score 52; DB 8; Length 174034;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-16;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCGAGTGTGATCAGCTGAGGCGAGAGTGTGAGAGCGAGCTGAGCAACAT 2940  
 |||||  
 Db 61038 GAGGCGAGTGTGATCAGCTGAGGCGAGAGTGTGAGAGCGAGCTGAGCAACAT 61089

RESULT 154  
 AC127468/c  
 LOCUS  
 DEFINITION Papio anubis clone RP41-26378, WORKING DRAFT SEQUENCE, 15 ordered pieces.  
 AC127468  
 AC127468.4 GI:31415891  
 HTG: HTGS PHASE2: HTGS DRAFT.  
 SOURCE Papio anubis (olive baboon)  
 ORGANISM Papio anubis  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Cercopithecoidea; Cercopithecoidea; Papio.  
 1 (bases 1 to 174521)  
 Antonellis, A., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakeley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S., Carliaga, K., Chu, G., Coleman, B., Coleman, H., Engle, J., Granite, S., Guan, X., Gupta, J., Haghighi, P., Han, J., Hansen, N., Ho, S.-L., Hu, P., Hurle, B., Idol, J.R., Karlins, B., Kwong, P., Laric, P., Lee-Jin, S.-O., Legaspi, R., Meduro, O.L., Maduro, V.B., Margulies, E.H., Masiello, C., Mascheri, B., McDowell, J., Pasutigian, C., Pearson, R., Portnoy, M.E., Prasad, A., Reddix-Dugue, N., Schandler, K., Schneider, P.G., Shan, K., Sison, C., Stantipod, S., Thomas, J.W., Thomas, P.J., Tsipouni, V., Vogt, J.L., Wetherby, K.D., Wiggins, L., Young, A. and Green, E.D.  
 NISC Comparative Sequencing Initiative  
 Unpublished  
 2 (bases 1 to 174521)  
 Green, E.D.  
 Direct Submission  
 Submitted (17-JUL-2002) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA  
 REFERENCE 3 (bases 1 to 174521)  
 Green, E.D.  
 Direct Submission  
 Submitted (05-JUN-2003) NIH Intramural Sequencing Center, 8717 Grovemont Circle, Gaithersburg, MD 20877, USA  
 COMMENT On Jun 5, 2003 this sequence version replaced gi:26050953.

----- Genome Center  
 Center: NIH Intramural Sequencing Center  
 Center code: NISC  
 Web site: <http://www.nisc.nih.gov>  
 Contact: [nisc\\_zoemhgrt.nih.gov](mailto:nisc_zoemhgrt.nih.gov)  
 ----- Project Information  
 Center project name: cyr  
 Center clone name: 263F08

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics  
 Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 171309 bases at least Q40  
 Consensus quality: 172311 bases at least Q30  
 Consensus quality: 172845 bases at least Q20  
 Insert size: 177000; agarose-fp  
 Insert size: 173121; sum-of-contigs  
 Quality coverage: 10.09x in Q20 bases; agarose-fp  
 Quality coverage: 10.32x in Q20 bases; sum-of-contigs

-----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 15 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.

1 42198: contig of 42198 bp in length  
 \* 42199 42298: gap of unknown length  
 \* 42299 45586: contig of 3388 bp in length  
 \* 45587 45786: gap of unknown length  
 \* 45787 71568: contig of 25782 bp in length  
 \* 71569 71668: gap of unknown length  
 \* 71669 75986: contig of 4318 bp in length  
 \* 75987 76086: gap of unknown length  
 \* 76087 85283: contig of 9197 bp in length  
 \* 85284 85383: gap of unknown length  
 \* 85384 103813: contig of 18430 bp in length  
 \* 103814 103913: gap of unknown length  
 \* 103914 105217: contig of 1304 bp in length  
 \* 105218 105317: gap of unknown length  
 \* 105318 112563: contig of 7246 bp in length  
 \* 112564 112663: gap of unknown length  
 \* 112664 114712: contig of 2049 bp in length  
 \* 114713 114812: gap of unknown length  
 \* 114813 141386: contig of 26574 bp in length  
 \* 141387 141486: gap of unknown length  
 \* 141487 144489: contig of 2963 bp in length  
 \* 144490 144549: gap of unknown length  
 \* 144550 148916: contig of 4367 bp in length  
 \* 148917 149016: gap of unknown length  
 \* 149017 153347: contig of 4231 bp in length  
 \* 153248 153347: gap of unknown length  
 \* 153348 165312: contig of 11865 bp in length  
 \* 165313 165312: gap of unknown length  
 \* 165313 174521: contig of 9209 bp in length.  
 Location/Qualifiers  
 1. 174521

FEATURES  
 source

```
/organism="Papio anubis"
/mol_type="genomic DNA"
/db_xref="taxon:9555"
/clone="RP41-263F8"
/clone_1lb="RP41"
1..42198
misc_feature
  /note="assembly fragment
  missing T7 clone end on 5' end of insert"
gap
  42199..42298
misc_feature
  /estimated_length=unknown
  42299..45686
  /note="assembly fragment"
gap
  45687..45786
  /estimated_length=unknown
  45787..71568
  /note="assembly fragment"
gap
  71569..71668
  /estimated_length=unknown
  71669..75986
  /note="assembly fragment"
gap
  75987..76086
  /estimated_length=unknown
  76087..85283
  /note="assembly fragment"
gap
  85284..85383
  /estimated_length=unknown
  85384..103813
  /note="assembly fragment"
gap
  103814..103913
  /estimated_length=unknown
  103914..105217
  /note="assembly fragment"
gap
  105218..105317
  /estimated_length=unknown
  105318..174521
  /note="clone overlaps with GenBank Accession Number
  AC127469 clone RP41-91123 (center project name cyq)"
misc_feature
  105318..112563
  /note="assembly fragment"
gap
  112564..112663
  /estimated_length=unknown
  112664..114712
  /note="assembly fragment"
gap
  114713..114812
  /estimated_length=unknown
  114813..141386
  /note="assembly fragment"
gap
  141387..141486
  /estimated_length=unknown
  141487..144449
  /note="assembly fragment"
gap
  144450..144549
  /estimated_length=unknown
  144550..148916
  /note="assembly fragment"
gap
  148917..149016
  /estimated_length=unknown
  149017..153247
  /note="assembly fragment"
gap
  153248..153347
  /estimated_length=unknown
  153348..165212
  /note="assembly fragment"
gap
  165213..165312
  /estimated_length=unknown
  165313..174521
  /note="assembly fragment
  clone end:SP6
  vector_side:right"
ORIGIN
```

```
Query Match 1.7%; Score 52; DB 14; Length 174521;
Best Local Similarity 100.0%; Pred. No. 5.7e-16;
```

```
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 2889 GAGGAGGTGATCCTGAGGCCGAGATTGAAACCACTGGCCCAACAT 2940
Db 90300 GAGGACGTGATCACTGAGGCCGAGATTGAAACCACTGGCCCAACAT 90249

RESULT 155
CONS01DX3/c 176343 bp DNA linear PRI 26-APR-2001
LOCUS Human chromosome 14 DNA sequence BAC R-517013 of library RPCT-11
DEFINITION from chromosome 14 of Homo sapiens (human), complete sequence.
ACCESSION AL139021
VERSION AL139021.6 GI:13159194
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
REFERENCE 1 (bases 1 to 176343)
AUTHORS Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,
Brottier,P., Catolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,
Levy,M., Bckenberg,R., Bruls,T., deBardindis,V., Cruaud,C.,
Gyapay,G., Saurin,W. and Weissenbach,J.
Sequencing of the human chromosome 14
Unpublished
2 (bases 1 to 176343)
Genoscope.
Direct Submission
Submitted (26-APR-2001) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
On Feb 27, 2001 this sequence version replaced gi:1161151.
----- Genome Center
Center: Genoscope / Centre National de Sequencage
Genetic code: GS
Web site: http://www.genoscope.cns.fr/
Contact: segref@genoscope.cns.fr
-----
The following BAC sequence is oriented from the T7 to the SP6 end.
Upstream BAC (overlapping the T7 end) : C-2002H8
Downstream BAC (overlapping the SP6 end) : R-401H9 (AC=AL135752)
----- Summary Statistics
Assembly program: Phrap, version 2.0
Quality coverage: 8.11x in Q20 bases; sum-of-contigs
-----
Overall quality chart :
Range : bases
0 - 9 :
1 - 9 : 2
10 - 19 : 10
20 - 29 : 210
30 - 39 : 4379
40 - 49 : 10906
50 - 59 : 9685
60 - 69 : 18251
70 - 79 : 45109
80 - 89 : 87791
90 - 99 :
Percentage of bases with a quality value >= 40 : 99 %.
-----
Location/Qualifiers
1..176343
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="14"
/clone="R-517013"
/clone_1lb="RPCT-11"
2338..2477
/note="matching EMBL:G33047"
FEATURES
source
STS
```

STS RHdb:RH67742  
dbSTS:STS47669  
Identified using the e-PCR software (G. Schuler) "

STS 35745. .35866  
/note="matching EMBL:G13592  
RHdb:RH5857  
RHdb:RH13847  
dbSTS:STS13237  
Identified using the e-PCR software (G. Schuler) "

STS 37127. .37251  
/note="matching EMBL:G30370  
RHdb:RH53945  
RHdb:RH38563  
dbSTS:STS8453  
Identified using the e-PCR software (G. Schuler) "

STS 66498. .66635  
/note="matching EMBL:N55165  
RHdb:RH66721  
dbSTS:STS46650  
Identified using the e-PCR software (G. Schuler) "

STS 71248. .71368  
/note="matching EMBL:H97583  
RHdb:RH78893  
dbSTS:STS55634  
Identified using the e-PCR software (G. Schuler) "

STS 72428. .72563  
/note="matching EMBL:N22461  
RHdb:RH46962  
dbSTS:STS40024  
Identified using the e-PCR software (G. Schuler) "

STS 110431. .110732  
/note="matching EMBL:G03433  
RHdb:RH34713  
dbSTS:STS1533  
Identified using the e-PCR software (G. Schuler) "

ORIGIN

Query Match 1.7% Score 52; DB 8; Length 176343;  
Best Local Similarity 100.0%; Pred. No. 5.7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTCACCTGCTCCAGCTGGCAACAGAGCAAGACTCTGTCTC 3122  
|||||  
4072 CAAGTTTGCCACCTGCACTCCAGCTGGCAACAGAGCAAGACTCTGTCTC 4021  
|||||

Db 4072 CAAGTTTGCCACCTGCACTCCAGCTGGCAACAGAGCAAGACTCTGTCTC 4021

RESULT 156  
AC150824 176871 bp DNA linear HTG 16-SEP-2004  
LOCUS Callitrix jacchus clone CH259-392G13, WORKING DRAFT SEQUENCE, 6  
DEFINITION Ordered pieces.  
AC150824  
AC150824  
VERSION AC150824.2 GI:52138777  
KEYWORDS HTG; HTGS PHASE2; HTGS DRAFT.  
SOURCE Callitrix jacchus (white-eufted-ear marmoset)  
ORGANISM Callitrix jacchus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Platyrrhini;  
Callitricidae; Callitrix.  
1 (bases 1 to 176871)  
Antenellia,A., Ayele,K., Benjamin,B., Blakesley,R.W., Boake,A.,  
Bouffard,G.G., Brinkley,C., Brooke,S., Chu,G., Coleman,H., Daki,N.,  
Engle,J., Guan,X., Gupta,J., Haghighi,P., Han,U., Hansen,N.,  
Ho,S.-L., Hu,P., Hurle,B., Idol,J.R., Jones,C., Kwong,P., Latic,P.,  
Larson,S., Lee-Lin,S.-O., Legaapi,R., Madden,M., Maduro,Q.L.,  
Maduro,V.B., Margulies,E.H., Masello,C., Maskeri,B., McDowell,J.,  
Mullikin,J.C., Park,M., Portnoy,M.E., Prasad,A., Puri,O.,  
Reditz-Dugue,N., Roasas,B., Schandler,K., Schueler,M.G., Shah,K.,  
Sison,C., Stancipop,S., Stephen,E., Thomas,J.W., Thomas,P.J.,  
Tajpouri,V., Vogt,J.U., Wetherby,K.D., Young,A. and Green,E.D.  
NISC Comparative Sequencing Initiative  
Unpublished  
2 (bases 1 to 176871)

AUTHORS Green, E.D.  
TITLE Direct Submission  
JOURNAL Submitted (11-AUG-2004) NIH Intramural Sequencing Center, 8717  
Government Circle, Gaithersburg, MD 20877, USA  
3 (bases 1 to 176871)  
REFERENCES Green, E.D.  
TITLE Direct Submission  
JOURNAL Submitted (16-SEP-2004) NIH Intramural Sequencing Center, 8717  
Government Circle, Gaithersburg, MD 20877, USA  
On Sep 16, 2004 this sequence version replaced gi:51101028.  
----- Genome Center  
Center: NIH Intramural Sequencing Center  
Center code: NISC  
Web site: <http://www.nisc.nih.gov>  
Contact: nisc.zoo@nih.gov  
----- Project Information  
Center project name: h3b  
Center clone name: 392G13

The sequence data in this record represents an 'enhanced' version of a Phase 2 submission. Specifically, the indicated order and orientation of each sequence contig has been established using one or more of the following: read-pair data from individual subclones, overlaps with neighboring clones, alignment with available reference sequence (e.g., human), and/or confirmation by PCR testing. In addition, the sequence assembly is based on at least 8X average coverage in Q20 bases and has been reviewed to rule out gross misassemblies, the low-quality ends of sequence contigs have been trimmed away, and each base is associated with a Phrap-derived quality score.

----- Summary Statistics  
Sequencing vector: plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye, 100% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 175488 bases at least Q40  
Consensus quality: 176005 bases at least Q30  
Consensus quality: 176238 bases at least Q20  
Insert size: 225000; agarose-gel  
Insert size: 176371; sum-of-contigs  
Quality coverage: 8.55x in Q20 bases; agarose-gel  
Quality coverage: 10.91x in Q20 bases; sum-of-contigs

-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 6 contigs. Gaps between the contigs  
\* are represented as runs of N. The order of the pieces  
\* is believed to be correct as given, however the sizes  
\* of the gaps between them are based on estimates that have  
\* provided by the submitter.  
\* This sequence will be replaced  
\* by the finished sequence as soon as it is available and  
\* the accession number will be preserved.  
\* 1 22851: contig of 22851 bp in length  
\* 22852 22951: gap of unknown length  
\* 22952 25041: contig of 2090 bp in length  
\* 25042 25141: gap of unknown length  
\* 25142 63108: contig of 37967 bp in length  
\* 63109 63208: gap of unknown length  
\* 63209 156929: contig of 93721 bp in length  
\* 156930 157029: gap of unknown length  
\* 157030 171090: contig of 14061 bp in length  
\* 171091 171191: gap of unknown length  
\* 171191 176871: contig of 5681 bp in length.  
Location/Qualifiers  
1. 176871  
/organism="Callitrix jacchus"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9483"  
/clone="CH259-392G13"  
/clone\_11b="CH259"  
/note="BAC resource: <http://bacpac.choi.org/>"  
1. 106470  
/note="clone overlaps with GenBank Accession Number

```
misc_feature      AC151042 clone CH259-63D17 (center project name hja)
1. 22851
   /note="assembly_fragment"
   clone_end:17
   vector_side:left"
gap              22852..22951
               /estimated_length=unknown
misc_feature      22952..25041
               /note="assembly_fragment"
gap              25042..25141
               /estimated_length=unknown
misc_feature      25142..63108
               /note="assembly_fragment"
gap              63109..63208
               /estimated_length=unknown
misc_feature      63209..156929
               /note="assembly_fragment"
misc_feature      79212..176871
               /note="clone overlaps with GenBank Accession Number
AC151035 clone CH259-434P22 (center project name hjc)"
gap              156930..157029
               /estimated_length=unknown
misc_feature      157030..171090
               /note="assembly_fragment"
gap              171091..171190
               /estimated_length=unknown
misc_feature      171191..176871
               /note="assembly_fragment"
               clone_end:SP6
               vector_side:right"

ORIGIN
Query Match      1.7%: Score 52; DB 14; Length 176871;
Best Local Similarity 100.0%; Pred. No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy              3071 CAAGATTGTGCACCTGCACTCCAGCTCGGCAACAGAGCAAGACTCTGTCTC 3122
Db              4700 CAAGATTGTGCACCTGCACTCCAGCTCGGCAACAGAGCAAGACTCTGTCTC 4751

RESULT 157
AC073954/c      177744 bp      DNA      linear      HTG 17-NOV-2000
LOCUS           Homo sapiens chromosome 17 clone RP11-147C10, WORKING DRAFT
DEFINITION      SEQUENCE, 34 unordered pieces.
ACCESSION       AC073954
VERSION         AC073954.3 GI:10998993
KEYWORDS        HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE          Homo sapiens (human)
ORGANISM        Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE        1 (bases 1 to 177744)
AUTHORS         Waterston,R.H.
TITLE           The sequence of Homo sapiens clone
JOURNAL          Unpublished
REFERENCE        2 (bases 1 to 177744)
AUTHORS         Waterston,R.H.
TITLE           Direct Substitution
JOURNAL          Submitted (07-JUL-2000) Genome Sequencing Center, Washington
MO 63108, USA   University School of Medicine, 4444 Forest Park Parkway, St. Louis,
On Oct 25, 2000 this sequence version replaced gi:9186849.

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0147C10
----- Summary Statistics -----
```

```
Sequencing vector: M13; 100%
Sequencing vector: plasmid; 0%
Chemistry: Dye-primer RT; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 155679 bases at least Q40
Consensus quality: 163329 bases at least Q30
Consensus quality: 166780 bases at least Q20
Insert size: 185000; agarose-fp
Insert size: 183559; sum-of-contigs
Quality coverage: 3.29 in Q20 bases; agarose-fp
Quality coverage: 3.43 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 34 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1      1247: contig of 1247 bp in length
*      1248      1347: gap of unknown length
*      1348      2762: contig of 1415 bp in length
*      2763      2862: gap of unknown length
*      2863      4773: contig of 1911 bp in length
*      4774      4873: gap of unknown length
*      4874      6326: contig of 1453 bp in length
*      6327      6426: gap of unknown length
*      6427      8873: contig of 2447 bp in length
*      8874      10783: gap of unknown length
*      10784      10883: gap of 1810 bp in length
*      10884      12763: gap of unknown length
*      12764      12863: gap of unknown length
*      12864      15381: contig of 2518 bp in length
*      15382      17365: contig of 1884 bp in length
*      17366      17465: gap of unknown length
*      17466      19992: contig of 2527 bp in length
*      19993      20092: gap of unknown length
*      20093      23512: contig of 3420 bp in length
*      23513      25699: gap of unknown length
*      25699      25799: contig of 2087 bp in length
*      25799      25799: gap of unknown length
*      25800      28333: contig of 2534 bp in length
*      28334      28434      32027: contig of 3594 bp in length
*      32028      32127: gap of unknown length
*      32128      36115: contig of 3988 bp in length
*      36116      36215: gap of unknown length
*      36216      39882: contig of 3667 bp in length
*      39883      39982: gap of unknown length
*      39983      44009: contig of 4027 bp in length
*      44010      44109: gap of unknown length
*      44110      47829: contig of 3720 bp in length
*      47830      47929: gap of unknown length
*      47930      53768: contig of 5839 bp in length
*      53769      53868: gap of unknown length
*      53869      60215: contig of 6347 bp in length
*      60216      60315: gap of unknown length
*      60316      66477: contig of 6162 bp in length
*      66478      66577: gap of unknown length
*      66578      71476: contig of 4899 bp in length
*      71477      71576: gap of unknown length
*      71577      78883: contig of 7307 bp in length
*      78884      78983: gap of unknown length
*      78984      85952: contig of 6369 bp in length
*      85953      86052: gap of unknown length
*      86053      93510: contig of 7458 bp in length
*      93511      93610: gap of unknown length
*      93611      100718: contig of 7108 bp in length
*      100719      100818: gap of unknown length
```

	*	100819	109795: contig of 8977 bp in length
	*	109796	109895: gap of unknown length
	*	109896	117790: contig of 7895 bp in length
	*	117791	126044: gap of unknown length
	*	117891	126044: contig of 8154 bp in length
	*	126045	126144: gap of unknown length
	*	126145	134442: contig of 828 bp in length
	*	134443	134442: gap of unknown length
	*	134543	145663: contig of 1121 bp in length
	*	145664	145789: gap of unknown length
	*	145764	157289: contig of 11426 bp in length
	*	157190	157289: gap of unknown length
	*	157290	166822: contig of 9533 bp in length
	*	166823	166922: gap of unknown length
	*	166923	177744: contig of 10822 bp in length
FEATURES			Location/Qualifiers
source			1..177744
			/organism="Homo sapiens"
			/mol_type="genomic DNA"
			/db_xref="taxon:9606"
			/chromosome="17"
			/clone="RP11-147C10"
		1..1247	
misc_feature		/note="assembly_name:Contig26"	
gap		1248..1347	
misc_feature		/estimated_length=unknown	
gap		1348..2762	
misc_feature		/note="assembly_name:Contig31"	
gap		2763..2862	
misc_feature		/estimated_length=unknown	
gap		2863..4773	
misc_feature		/note="assembly_name:Contig34"	
gap		4774..4873	
misc_feature		/estimated_length=unknown	
gap		4874..6326	
misc_feature		/note="assembly_name:Contig35"	
gap		6327..6426	
misc_feature		/estimated_length=unknown	
gap		6427..8873	
misc_feature		/note="assembly_name:Contig36"	
gap		8874..8973	
misc_feature		/estimated_length=unknown	
gap		8974..10783	
misc_feature		/note="assembly_name:Contig37"	
gap		10784..10883	
misc_feature		/estimated_length=unknown	
gap		10884..12763	
misc_feature		/note="assembly_name:Contig38"	
gap		12764..12863	
misc_feature		/estimated_length=unknown	
gap		12864..15381	
misc_feature		/note="assembly_name:Contig39"	
gap		15382..15481	
misc_feature		/estimated_length=unknown	
gap		15482..17365	
misc_feature		/note="assembly_name:Contig40"	
gap		17366..17465	
misc_feature		/estimated_length=unknown	
gap		17466..19592	
misc_feature		/note="assembly_name:Contig41"	
gap		19593..20092	
misc_feature		/estimated_length=unknown	
gap		20093..23512	
misc_feature		/note="assembly_name:Contig42"	
gap		23513..23612	
misc_feature		/estimated_length=unknown	
gap		23613..25699	
misc_feature		/note="assembly_name:Contig43"	
gap		25700..25799	
misc_feature		/estimated_length=unknown	
gap		25800..28333	
misc_feature		/note="assembly_name:Contig44"	
gap		28334..28433	

Query Match	1.7%; Score 52; DB 14; Length 177744;
Best Local Similarity	100.0%; Pred. No. 5,7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Db	146170 GAGCAGATGGATCACTGAGGCCAGAGTTCCAGACACACCTGGCCACAT 146119
Result 158	
LOCUS	AC010761 177773 bp DNA linear PRI 10-JAN-2003
DEFINITION	Homo sapiens chromosome 17, clone RP11-386F9, complete sequence.
ACCESSION	AC010761
VERSION	AC010761.10 GI:25753488
KEYWORDS	HTG.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 177773)
AUTHORS	Bitnen,B., Nussbaum,C., Lander,B., Allen,N., Anderson,S., Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeRellano,K., Dewar,K., Domingo,M., Donelan,L., Doyle,M., Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D., Hagland,J., Gardina,S., Grant,G., Hagoeb,B., Headford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karalas,A., Klein,J., Lehoccky,J., Lien,C., Locke,K., MacDonald,P., Marquis,N., McMan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Teefaye,S., Tirrell,A., Vassiliou,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.Y., Zimmer,A. and Zody,M.
REFERENCE	2 (bases 1 to 177773)
AUTHORS	Bitnen,B., Linton,L., Nussbaum,C., Lander,B., Allen,N., Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Boguslavsky,L., Boughalter,B., Brown,A., Castle,A., Colangelo,M., Collins,S., Collymore,A., Cooke,P., DeRellano,K., Dewar,K., Domingo,M., Donelan,L., Doyle,M., Ferreira,P., Fitzhugh,W., Forrest,C., Funke,R., Gage,D., Hagland,J., Gardina,S., Grant,G., Hagoeb,B., Headford,A., Horton,L., Howland,J.C., Johnson,R., Jones,C., Kann,L., Karalas,A., Klein,J., Lehoccky,J., Lien,C., Locke,K., MacDonald,P., Marquis,N., McMan,P., McGurk,A., McKernan,K., McLaughlin,J., Meldrum,J., Morrow,J., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P., Peterson,K., Pollara,V., Riley,R., Roy,A., Santos,R., Severy,P., Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J., Teefaye,S., Tirrell,A., Vassiliou,H., Vo,A., Wheeler,J., Wu,X., Wyman,D., Ye,W.Y., Zimmer,A. and Zody,M.
REFERENCE	3 (bases 1 to 177773)
AUTHORS	Bitnen,B., Nussbaum,C., Lander,B., Ali,A., Allen,N., Anderson,S., Barna,N., Baetjen,V., Bloom,T., Boguslavsky,L., Boughalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cook,A., Cooke,P., DeRellano,K., Dewar,K., Diaz,J.S., Dodge,S.,



Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N., Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Mathews, C., McCarthy, M., Meldrum, J., McDonald, L., Mhova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhthang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testafaye, S., Theodore, J., Tophan, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (11-NOV-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 177773)

TITLE  
JOURNAL  
REFERENCE  
AUTHORS

Barra, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, X., Collymore, A., Cook, A., Cooke, P., Dearellano, R., Dewar, K., Diaz, J.S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N., Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Mathews, C., McCarthy, M., Meldrum, J., McDonald, L., Mhova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunhthang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testafaye, S., Theodore, J., Tophan, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

Direct Submission  
Submitted (10-JAN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jan 10, 2003 this sequence version replaced gi:24850532.  
All repeats were identified using RepeatMasker:  
Smith, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

TITLE  
JOURNAL  
COMMENT

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: MIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: 11611  
Center clone name: 386\_F\_9

FRATRES  
source  
-----  
Location/Qualifiers  
1. 177773  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/map="17"  
/clone="RP11-386F9"  
/clone\_1fb="RP11-386F9 Human Male BAC"  
31. .91  
repeat\_region  
/rpt\_family="MIR"  
92. .396  
repeat\_region  
/rpt\_family="AluY"  
397. .443  
repeat\_region  
/rpt\_family="MIR"  
1541. .1662  
repeat\_region  
/rpt\_family="AluYb"  
1663. .1915  
repeat\_region  
/rpt\_family="AluSp"  
1916. .2068  
repeat\_region  
/rpt\_family="AluYb"

repeat\_region complement(2094. .2403)  
/rpt\_family="AluSp"  
repeat\_region complement(2621. .2655)  
/rpt\_family="PAM"  
2656. .2983  
repeat\_region complement(2984. .3026)  
/rpt\_family="PAM"  
3027. .3215  
repeat\_region complement(3230. .3515)  
/rpt\_family="AluSp"  
3607. .3773  
repeat\_region  
/rpt\_family="MIR"  
3932. .3953  
repeat\_region  
/rpt\_family="(TG)n"  
3956. .4172  
repeat\_region  
/rpt\_family="L1MC3"  
4177. .4349  
repeat\_region  
/rpt\_family="PAM"  
complement(4456. .4759)  
/rpt\_family="AluSp"  
4772. .5682  
repeat\_region  
/rpt\_family="L1MC3"  
5683. .6006  
repeat\_region  
/rpt\_family="AluSp"  
6007. .6108  
repeat\_region  
/rpt\_family="L1MC3"  
6181. .6455  
repeat\_region  
/rpt\_family="AluYb"  
6685. .6821  
repeat\_region  
/rpt\_family="AluSp"  
6822. .7132  
repeat\_region  
/rpt\_family="AluSp"  
7133. .7294  
repeat\_region  
/rpt\_family="AluSp"  
8303. .8367  
repeat\_region  
/rpt\_family="Alu"  
complement(8373. .8421)  
/rpt\_family="Alu"  
8431. .8455  
repeat\_region  
/rpt\_family="AT\_rich"  
8456. .8759  
repeat\_region  
/rpt\_family="AluSp"  
9104. .9406  
repeat\_region  
/rpt\_family="AluSp"  
9419. .9536  
repeat\_region  
/rpt\_family="PAM\_C"  
complement(9557. .9849)  
/rpt\_family="AluYb"  
9850. .9883  
repeat\_region  
/rpt\_family="AT\_rich"  
9886. .10131  
repeat\_region  
/rpt\_family="AluYb"  
complement(10405. .11595)  
/rpt\_family="SVA"  
complement(11607. .12423)  
/rpt\_family="SVA"  
complement(12650. .12763)  
/rpt\_family="MER5A"  
13119. .13194  
repeat\_region  
/rpt\_family="L2"  
13236. .13255  
repeat\_region  
/rpt\_family="(A)n"  
13491. .13699  
repeat\_region  
/rpt\_family="MIR"  
complement(15981. .16105)  
/rpt\_family="PAM"  
16137. .16377  
repeat\_region  
/rpt\_family="L2"  
16378. .16660  
repeat\_region  
/rpt\_family="AluSp"  
16661. .17274  
repeat\_region

```

repeat_region      /rpt_family="L2"
                    17513..17625
repeat_region      /rpt_family="L2"
                    17629..17922
repeat_region      /rpt_family="AluSx"
                    17924..17949
repeat_region      /rpt_family=" (TATATG) n"
                    17950..18079
repeat_region      /rpt_family="FRAM"
                    18097..18145
repeat_region      /rpt_family="L2"
                    complement(18703..19008)
repeat_region      /rpt_family="AluSx"
                    19259..19546
repeat_region      /rpt_family="AluY"
                    19538..19586
repeat_region      /rpt_family=" (TA) n"

```

Query Match 1.7%; Score 52; DB 8; Length 177773;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-16;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
 Db 45101 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 45152

RESULT 159  
 AC012350/c 178168 bp DNA linear HTG 01-APR-2000  
 DEFINITION Homo sapiens clone RP11-16N9, WORKING DRAFT SEQUENCE, 16 unordered  
 pieces

AC012350  
 VERSION AC012350.3 GI:7381803  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 1 (bases 1 to 178168)  
 Birren, B., Linton, L., Nusbaum, C. and Lander, B.  
 Homo sapiens, clone RP11-16N9  
 Unpublished  
 2 (bases 1 to 178168)

Birren, B., Linton, L., Nusbaum, C., Lander, B., Allen, N., Anderson, M.,  
 Baldwin, J., Batra, N., Beckerly, R., Boguslavsky, L., Boukhgalter, B.,  
 Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A.,  
 Cooke, P., Dearellano, K., Dewar, K., Domingo, M., Donelan, L., Doyle, M.,  
 Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D.,  
 Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heath, A., Horton, L.,  
 Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatsis, A., Klein, J.,  
 Lehoczy, J., Lien, C., Locke, K., MacDonald, P., Margulis, N.,  
 McGowan, P., McGurk, A., McKernan, K., McDonald, J., Meldrum, J.,  
 Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P.,  
 Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P.,  
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
 Testa, S., Tittell, A., Vassiliev, H., Vo, A., Wheeler, J., Wu, X.,  
 Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

Direct Submission  
 Submitted (25-OCT-1999) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Apr 1, 2000 this sequence version replaced gi:6479001.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L3593

Center clone name: 16 N 9  
 ----- Summary Statistics  
 Sequencing vector: M13; M77815, 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.960731  
 Consensus quality: 172070 bases at least Q40  
 Consensus quality: 174479 bases at least Q30  
 Consensus quality: 175597 bases at least Q20  
 Insert size: 170000; agarose-gel  
 Insert size: 176668; sum-of-contigs  
 Quality coverage: 6.4 in Q20 bases; agarose-gel  
 Quality coverage: 6.2 in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 16 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.

```

1 1694: contig of 1694 bp in length
1 1695: gap of 100 bp
1 1795: contig of 2643 bp in length
1 4438: gap of 100 bp
1 4538: contig of 2107 bp in length
1 6645: gap of 100 bp
1 6745: contig of 3348 bp in length
1 10093: gap of 100 bp
1 10193: contig of 1609 bp in length
1 11802: gap of 100 bp
1 11902: contig of 3653 bp in length
1 15584: gap of 100 bp
1 15685: contig of 3951 bp in length
1 15685: gap of 100 bp
1 16775: contig of 8158 bp in length
1 19776: gap of 100 bp
1 27914: contig of 6143 bp in length
1 28014: gap of 100 bp
1 34157: contig of 9069 bp in length
1 34257: gap of 100 bp
1 43326: contig of 11840 bp in length
1 43426: gap of 100 bp
1 55265: contig of 12941 bp in length
1 55365: gap of 100 bp
1 55365: contig of 12941 bp in length
1 68307: gap of 100 bp
1 68406: contig of 14504 bp in length
1 82910: gap of 100 bp
1 82911: contig of 19171 bp in length
1 83011: gap of 100 bp
1 102182: contig of 25777 bp in length
1 102282: gap of 100 bp
1 128058: contig of 50010 bp in length
1 128158: gap of 100 bp
1 178159: contig of 50010 bp in length
1 178168: Location/Qualifiers
1 178168
  /organism="Homo sapiens"
  /mol_type="genomic DNA"
  /db_xref="taxon:9606"
  /clone_id="RP11-16N9"
  /clone_lib="RP11-16N9 Human Male BAC"
1 1694
  /note="assembly_fragment"
1695..1794
  /estimated_length=100
1795..4437
  /note="assembly_fragment"
4438..4537
  /estimated_length=100
4538..6644
  /note="assembly_fragment"
6645..6744
  /estimated_length=100
6745..10092
  /note="assembly_fragment"

```

misc\_feature 1..1694  
 /note="assembly\_fragment"  
 gap 1695..1794  
 /estimated\_length=100  
 misc\_feature 1795..4437  
 /note="assembly\_fragment"  
 gap 4438..4537  
 /estimated\_length=100  
 misc\_feature 4538..6644  
 /note="assembly\_fragment"  
 gap 6645..6744  
 /estimated\_length=100  
 misc\_feature 6745..10092  
 /note="assembly\_fragment"

```

gap          10093..10192
             /estimated_length=100
misc_feature 10193..11801
             /note="assembly_fragment
             /note_end:SP6
             vector_side:right"
gap          11802..11901
             /estimated_length=100
misc_feature 11902..15584
             /note="assembly_fragment"
gap          15585..15684
             /estimated_length=100
misc_feature 15685..19675
             /note="assembly_fragment"
gap          19676..19775
             /estimated_length=100
misc_feature 19776..27913
             /note="assembly_fragment"
gap          27914..28013
             /estimated_length=100
misc_feature 28014..34156
             /note="assembly_fragment"
gap          34157..34256
             /estimated_length=100
misc_feature 34257..43325
             /note="assembly_fragment"
gap          43326..43425
             /estimated_length=100
misc_feature 43426..55265
             /note="assembly_fragment"
gap          55266..55365
             /estimated_length=100
misc_feature 55366..68306
             /note="assembly_fragment"
gap          68307..68406
             /estimated_length=100
misc_feature 68407..82510
             /note="assembly_fragment"
gap          82511..83010
             /estimated_length=100
misc_feature 83011..102181
             /note="assembly_fragment
             /note_end:T7
             vector_side:right"
gap          102182..102281
             /estimated_length=100
misc_feature 102282..128058
             /note="assembly_fragment"
gap          128059..128158
             /estimated_length=100
misc_feature 128159..178168
             /note="assembly_fragment"

ORIGIN
Query Match      1.7%; Score 52; DB 14; Length 178168;
Best Local Similarity 100.0%; Pred. No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3071 CAAGATTGTGCCACTGCACCTGCAGCTGGCAACAGAGCAAGACTGTGTCTC 3122
         |||||
Db       98028 CAAGATTGTGCCACTGCACCTGCAGCTGGCAACAGAGCAAGACTGTGTCTC 97977

RESULT 160
AC093126 181792 bp      DNA      linear      PRI 19-OCT-2002
LOCUS     AC093126/c
DEFINITION Papio anubis clone RP41-231J21, complete sequence.
ACCESSION AC093126
VERSION   AC093126.2  GI:24137422
KEYWORDS  HTG.
SOURCE    Papio anubis (olive baboon)
ORGANISM  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

REFERENCE
AUTHORS
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Cercopithecoidea; Cercopithecoinae; Papio.
1 (bases 1 to 181792)
Akhter,N., Antonellis,A., Ayele,K., Beckstrom-Stenberg,S.M.,
Benjamin,B., Blakesley,R.W., Bouffard,G.G., Brinkley,C., Brooks,S.,
Cariaga,K., Coleman,B., Dietrich,N.L., Engle,J., Granite,S.,
Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N., Ho,S.-L.,
Idol,J.R., Karlins,E., Loric,P., Lee-Lin,S.-O., Legaspi,R.,
Maduro,Q.L., Maduro,V.B., Margulies,E.H., Mariljo,C., Maskeri,B.,
Mastrion,S.D., McCloskey,J.C., McDowell,J., Pasutirgan,C.,
Pearson,R., Portnoy,M.B., Praad,A., Reddix-Dugue,N., Schandler,K.,
Schneider,M.G., Sison,C., Stancipop,S., Thomas,J.W., Thomas,P.J.,
Touchman,J.W., Vogt,J.L., Walker,M., Weherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.
MISC Comparative Sequencing Initiative
Unpublished
2 (bases 1 to 181792)
Green,E.D.
REFERENCE
AUTHORS
Direct Submission
Submitted (11-AUG-2001) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
3 (bases 1 to 181792)
Green,E.D.
REFERENCE
AUTHORS
Direct Submission
Submitted (19-OCT-2002) NIH Intramural Sequencing Center, 8717
Grovermont Circle, Gaithersburg, MD 20877, USA
On Oct 19, 2002 this sequence version replaced gi:15148116.
----- Genome Center
Center: NIH Intramural Sequencing Center
Center code: NISC
Web site: http://www.nisc.nih.gov
Contact: nisc.zoo@hgrl.nih.gov
----- Project Information
Center project name: ccu
Center clone name: 231J21

This sequence was finished as follows unless otherwise noted:
all regions were double-stranded, sequenced with an
alternate chemistry, or covered by high quality data
(i.e., phred quality >= 30); an attempt was made to resolve
all sequencing problems, such as compressions and repeats;
all regions were covered by at least one plasmid subclone
or more than one M13 subclone; and the assembly was confirmed
by restriction digest.

CLONE LENGTH: This sequence represents the entire insert of
this clone unless otherwise noted. If there are overlapping
clones, the overlaps are noted in the beginning and end of
the Features section.

FEATURES
source
location/Qualifiers
1..181792
/organism="Papio anubis"
/mol_type="genomic DNA"
/db_xref="taxon:9555"
/clone="RP41-231J21"
/clone_1lb="RP41"
19955..20051
/note="single clone coverage"
44150..44183
/note="single clone coverage"
55640..55715
/note="single clone coverage"
80993..81045
/note="single clone coverage"
158760..158799
/note="single clone coverage"

ORIGIN
Query Match      1.7%; Score 52; DB 8; Length 181792;
Best Local Similarity 100.0%; Pred. No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2889 GAGGAGGTGATCACTGAGGCGCAGGATTGAGACCAAGCTGGCCAAT 2940

```

Db 35625 GAGGCGAGTGGACCTGAGGCGCAGAGAGTTCAGAGCCAGCCCTGGCCACAT 35574

## RESULT 161

AC067929/c  
LOCUS AC067929/c  
DEFINITION Homo sapiens chromosome 6 clone RP11-403120 map 6, WORKING DRAFT

SEQUENCE, 18 unordered pieces.

AC067929  
AC067929.2 GI:8247824

VERSION HTG; HTGS PHASE1; HTGS\_DRAFT.

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

1 (bases 1 to 182152)

2 (bases 1 to 182152)

3 (bases 1 to 182152)

4 (bases 1 to 182152)

5 (bases 1 to 182152)

6 (bases 1 to 182152)

7 (bases 1 to 182152)

8 (bases 1 to 182152)

9 (bases 1 to 182152)

10 (bases 1 to 182152)

11 (bases 1 to 182152)

12 (bases 1 to 182152)

13 (bases 1 to 182152)

14 (bases 1 to 182152)

15 (bases 1 to 182152)

16 (bases 1 to 182152)

17 (bases 1 to 182152)

18 (bases 1 to 182152)

19 (bases 1 to 182152)

20 (bases 1 to 182152)

21 (bases 1 to 182152)

22 (bases 1 to 182152)

23 (bases 1 to 182152)

24 (bases 1 to 182152)

25 (bases 1 to 182152)

26 (bases 1 to 182152)

27 (bases 1 to 182152)

28 (bases 1 to 182152)

29 (bases 1 to 182152)

30 (bases 1 to 182152)

31 (bases 1 to 182152)

32 (bases 1 to 182152)

33 (bases 1 to 182152)

34 (bases 1 to 182152)

35 (bases 1 to 182152)

36 (bases 1 to 182152)

37 (bases 1 to 182152)

38 (bases 1 to 182152)

39 (bases 1 to 182152)

40 (bases 1 to 182152)

41 (bases 1 to 182152)

42 (bases 1 to 182152)

43 (bases 1 to 182152)

44 (bases 1 to 182152)

45 (bases 1 to 182152)

46 (bases 1 to 182152)

47 (bases 1 to 182152)

48 (bases 1 to 182152)

49 (bases 1 to 182152)

Smith, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/BM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIRB

Web site: http://www-seq.wi.mit.edu

Contact: sequence submissions@genome.wi.mit.edu

Project Information

Center project name: L10103

Center clone name: 403\_1\_20

Sequencing vector: M13; M77815; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 17395 bases at least Q40

Consensus quality: 179105 bases at least Q30

Insert size: 179000; agarose-fp

Insert size: 180452; sum-of-coverage

Quality coverage: 4.8 in Q20 bases; agarose-fp

Quality coverage: 4.7 in Q20 bases; sum-of-coverage

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 18 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

```

1      2222: contig of 2222 bp in length
2      2322: gap of 100 bp
3      2323: gap of 100 bp
4      3785: contig of 1463 bp in length
5      3786: gap of 100 bp
6      3885: gap of 100 bp
7      3886: contig of 2868 bp in length
8      6754: gap of 100 bp
9      6854: gap of 100 bp
10     10036: contig of 3183 bp in length
11     10137: gap of 100 bp
12     10137: contig of 4401 bp in length
13     14538: gap of 100 bp
14     14538: contig of 6540 bp in length
15     21177: gap of 100 bp
16     21177: contig of 8963 bp in length
17     21278: gap of 100 bp
18     30240: gap of 100 bp
19     30340: gap of 100 bp
20     30340: contig of 9783 bp in length
21     40123: gap of 100 bp
22     40124: gap of 100 bp
23     40223: gap of 100 bp
24     47999: contig of 7776 bp in length
25     48000: gap of 100 bp
26     48000: contig of 10136 bp in length
27     58335: gap of 100 bp
28     58335: contig of 8706 bp in length
29     67042: gap of 100 bp
30     67042: contig of 8706 bp in length
31     75921: gap of 100 bp
32     75921: contig of 8780 bp in length
33     75922: gap of 100 bp
34     75922: contig of 10506 bp in length
35     86527: gap of 100 bp
36     86527: contig of 14707 bp in length
37     101334: gap of 100 bp
38     101335: gap of 100 bp
39     101335: contig of 16196 bp in length
40     117631: gap of 100 bp
41     117631: contig of 20663 bp in length
42     138293: gap of 100 bp
43     138293: contig of 19448 bp in length
44     157841: gap of 100 bp
45     157842: gap of 100 bp
46     182152: contig of 24211 bp in length
47     157942: Location/Qualifiers
48     1.182152
49     /organism="Homo sapiens"
50     /mol_type="genomic DNA"
51     /db_xref="taxon:9606"
52     /chromosome="6"
53     /map="6"
54     /clone="RP11-403120"

```

## FEATURES

source

TITLE  
JOURNAL  
COMMENT  
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jun 4, 2000 this sequence version replaced gi:7655991.  
All repeats were identified using RepeatMasker:

```

misc_feature      /clone_11b="RPCT-11 Human Male BAC"
1..2222
/note="assembly_fragment"
gap              2223..2322
/estimated_length=100
misc_feature      2323..3785
/note="assembly_fragment"
gap              3786..3985
/estimated_length=100
misc_feature      3986..6753
/note="assembly_fragment"
gap              6754..6853
/estimated_length=100
misc_feature      6854..10036
/note="assembly_fragment"
gap              10037..10136
/estimated_length=100
misc_feature      10137..14537
/note="assembly_fragment"
gap              14538..14637
/estimated_length=100
misc_feature      14638..21177
/note="assembly_fragment"
gap              21178..21277
/estimated_length=100
misc_feature      21278..30240
/note="assembly_fragment"
gap              30241..30340
/estimated_length=100
misc_feature      30341..40123
/note="assembly_fragment"
gap              40124..40223
/estimated_length=100
misc_feature      40224..47939
/note="assembly_fragment"
clone_end:SP6
vector_side:left"
gap              48000..48039
/estimated_length=100
misc_feature      48100..58235
/note="assembly_fragment"
gap              58236..58335
/estimated_length=100
misc_feature      58336..67041
/note="assembly_fragment"
gap              67042..67141
/estimated_length=100
misc_feature      67142..75921
/note="assembly_fragment"
gap              75922..76021
/estimated_length=100
misc_feature      76022..86527
/note="assembly_fragment"
gap              86528..86627
/estimated_length=100
misc_feature      86628..101334
/note="assembly_fragment"
gap              101335..101434
/estimated_length=100
misc_feature      101435..117630

Query Match      1.7%; Score 52; DB 14; Length 182152;
Best Local Similarity 100.0%; Pred. No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY              3071 CAAGATTGCGCACTGCACCTCGGCAACAGAGCAAGACTGTCTC 3122
|||||
Db              14779 CAAGATTGCGCACTGCACCTCGGCAACAGAGCAAGACTGTCTC 14728
|||||

```

RESULT 162  
AC034244  
LOCUS

182892 bp

DNA

linear

PRI 20-JUL-2001

```

DEFINITION      Homo sapiens chromosome 5 clone RP11-101B14, complete sequence.
ACCESSION       AC034244
VERSION         AC034244.6
KEYWORDS        HTG.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE       1 (bases 1 to 182892)
AUTHORS        DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE          Direct Submission
JOURNAL        Unpublished
REFERENCE       2 (bases 1 to 182892)
AUTHORS        DOE Joint Genome Institute.
TITLE          Direct Submission
JOURNAL        Submitted (05-APR-2000) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE       3 (bases 1 to 182892)
AUTHORS        DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE          Direct Submission
JOURNAL        Submitted (26-JUN-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
REFERENCE       4 (bases 1 to 182892)
AUTHORS        DOE Joint Genome Institute and Stanford Human Genome Center.
TITLE          Direct Submission
JOURNAL        Submitted (20-JUL-2001) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT         On Jul 20, 2001 this sequence version replaced gi:14550305.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.shgc.stanford.edu
Quality: Phrap Quality >=40 99.8% of Sequence;
Estimated Total Number of Errors is 0.3.

FEATURES        source
                1..182892
                /organism="Homo sapiens"
                /mol_type="genomic DNA"
                /db_xref="taxon:9606"
                /chromosome="5"
                /clone="RP11-101B14"

ORIGIN

Query Match      1.7%; Score 52; DB 8; Length 182892;
Best Local Similarity 100.0%; Pred. No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY              3071 CAAGATTGCGCACTGCACCTCGGCAACAGAGCAAGACTGTCTC 3122
|||||
Db              91960 CAAGATTGCGCACTGCACCTCGGCAACAGAGCAAGACTGTCTC 92011
|||||

RESULT 163
AC019059/c      183556 bp      DNA      linear      HTG 07-JUL-2000
LOCUS           Homo sapiens chromosome 11 clone RP11-125P14, WORKING DRAFT
DEFINITION      SEQUENCE, 31 unordered pieces.
ACCESSION       AC019059
VERSION         AC019059.4
KEYWORDS        HTG; HTGS PHASE1; HTGS_DRAFT.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE       1 (bases 1 to 183556)
AUTHORS        Waterston,R.H.
TITLE          The sequence of Homo sapiens clone
JOURNAL        Unpublished
REFERENCE       2 (bases 1 to 183556)
AUTHORS        Waterston,R.H.
TITLE          Direct Submission

```





```

gap /note="assembly_name:Contig18"
33627..33726
misc_feature /estimated_length=unknown
33727..33741
gap /note="assembly_name:Contig19"
33742..33744
misc_feature /estimated_length=unknown
337442..41642
gap /note="assembly_name:Contig20"
41643..41742
gap /estimated_length=unknown
41743..47291
misc_feature /note="assembly_name:Contig21"
47292..47391
gap /estimated_length=unknown
47392..50826
misc_feature /note="assembly_name:Contig22"
50827..50926
gap /estimated_length=unknown
50927..58180
misc_feature /note="assembly_name:Contig23"
58181..58280
gap /estimated_length=unknown
58281..63165
misc_feature /note="assembly_name:Contig24"
63166..63265
gap /estimated_length=unknown
63266..68481
misc_feature /note="assembly_name:Contig25"
68482..68581
gap /estimated_length=unknown
68582..73841
misc_feature /note="assembly_name:Contig26"
73842..73941
gap /estimated_length=unknown
73942..79919
misc_feature

Query Match 1.7%; Score 52; DB 14; Length 183556;
Best Local Similarity 100.0%; Pred. No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGCGAGTGTGATCACCTGAGGCGAGAGTTCCAGACCAAGCTGCGCAACAT 2940
|||||
Db 127068 GAGCGAGTGTGATCACCTGAGGCGAGAGTTCCAGACCAAGCTGCGCAACAT 127017
|||||

RESULT 164 AC139565 184252 bp DNA linear HTG 20-FEB-2003
AC139565/c Homo sapiens chromosome 15 clone RP13-620J1 map 15, WORKING DRAFT
DEFINITION
SEQUENCE 4 unordered pieces.
ACCESSION AC139565
VERSION AC139565.2 GI:28436285
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 184252)
Birren,B., Nusbaum,C. and Lander,E.
Homo sapiens chromosome 15, clone RP13-620J1
REFERENCE JOURNAL
AUTHORS Unpublished
2 (bases 1 to 184252)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barina,N., Bastien,V., Bloom,T.,
Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Cornum,B., DeArelano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Halez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,

```

```

Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Melirim,J., Menus,L., Mihova,T., Mlangi,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Nordu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramsamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Strange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talmas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (05-FEB-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 184252)
Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
Anderson,S., Arachchi,H.M., Barina,N., Bastien,V., Bloom,T.,
Boguslavsky,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
Collymore,A., Cook,A., Cooke,P., Cornum,B., DeArelano,K.,
Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J., Gardyna,S.,
Graham,L., Grand-Pierre,N., Halez,N., Hagopian,D., Hagos,B.,
Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
Lindblad-Toh,K., Liu,G., Lui,A., Mabbitt,R., Maclean,C.,
Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
Melirim,J., Menus,L., Mihova,T., Mlangi,V., Murphy,T., Naylor,J.,
Nguyen,C., Nicol,R., Nordu,C., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
Rachupka,A., Ramsamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C.,
Spencer,B., Strange-Thomann,N., Stojanovic,N., Stubbs,M.,
Talmas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M.,
Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (20-FEB-2003) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Feb 20, 2003 this sequence version replaced gi:28209707.
All repeats were identified using RepeatMasker:
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: MIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L29407
Center clone name: 620 J.1
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960721
Consensus quality: 183540 bases at least Q40
Consensus quality: 183745 bases at least Q30
Consensus quality: 183810 bases at least Q20
Insert size: 181000; agarose-fp
Insert size: 183952; sum-of-ctnigs
Quality coverage: 18.0 in Q20 bases; agarose-fp
Quality coverage: 17.7 in Q20 bases; sum-of-ctnigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 89178: contig of 89178 bp in length
* 89179 89278: gap of 100 bp
* 89279 89403: contig of 125 bp in length

```

\* 89404 89503: gap of 100 bp  
\* 89504 90667: contig of 1164 bp in length  
\* 90668 90767: gap of 100 bp  
\* 90768 184252: contig of 93465 bp in length.

FEATURES  
source  
1.184252  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="15"  
/map="15"  
/clone="RP13-620J1"  
/clone\_1lb="RPC1-13 Human Female BAC"  
1.89178  
/note="assembly\_fragment  
clone\_end:SP6  
vector\_side:left"  
89179..89278  
/estimated\_length=100  
89279..89403  
/note="assembly\_fragment"  
89404..89503  
/estimated\_length=100  
89504..90667  
/note="assembly\_fragment"  
90668..90767  
/estimated\_length=100  
90768..184252  
/note="assembly\_fragment  
clone\_end:T7  
vector\_side:right"

ORIGIN  
Query Match 1.7%; Score 52; DB 14; Length 184252;  
Best Local Similarity 100.0%; Pred. No. 5.7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCGAGTGATCCTGAGGCGAGAGTTGAGAGCCAGCTCGGCCAACAT 2940  
DB 154414 GAGGCGAGTGATCCTGAGGCGAGAGTTGAGAGCCAGCTCGGCCAACAT 154363  
|||||

RESULT 165  
AC008675 186415 bp DNA linear PRI 20-DEC-2000  
LOCUS Homo sapiens chromosome 5 clone CTB-4518, complete sequence.  
AC008675  
AC008675.5 GI:11908270  
HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE  
AUTHORS 1 (bases 1 to 186415)  
TITLE DOE Joint Genome Institute and Stanford Human Genome Center.  
JOURNAL Direct Submission  
Unpublished  
2 (bases 1 to 186415)  
DOE Joint Genome Institute.  
REFERENCE  
AUTHORS Direct Submission  
TITLE Submitted (03-AUG-1999) Production Sequencing Facility, DOE Joint  
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 186415)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
REFERENCE  
AUTHORS Direct Submission  
TITLE Submitted (20-DEC-2000) DOE Joint Genome Institute, 2800 Mitchell  
JOURNAL Drive, Walnut Creek, CA 94598, USA  
On Dec 20, 2000 this sequence version replaced gi:7709256.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www-shgc.stanford.edu

Quality: Phrap Quality >=40 99.4% of Sequence;  
Estimated Total Number of Errors is 1.3.

FEATURES  
source  
1.186415  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="CTB-4518"

ORIGIN  
Query Match 1.7%; Score 52; DB 8; Length 186415;  
Best Local Similarity 100.0%; Pred. No. 5.7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GATCACCTGAGGCGAGAGTTGAGAGCCAGCTCGGCCAACATAGCAAAACC 2950  
DB 116437 GATCACCTGAGGCGAGAGTTGAGAGCCAGCTCGGCCAACATAGCAAAACC 116488  
|||||

RESULT 166  
AC018757/c 186418 bp DNA linear PRI 03-OCT-2001  
LOCUS Homo sapiens chromosome 5 clone CTB-4518, complete sequence.  
AC018757  
AC018757.6 GI:15887297  
HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE  
AUTHORS 1 (bases 1 to 186418)  
TITLE DOE Joint Genome Institute and Stanford Human Genome Center.  
JOURNAL Direct Submission  
Unpublished  
2 (bases 1 to 186418)  
DOE Joint Genome Institute.  
REFERENCE  
AUTHORS Direct Submission  
TITLE Submitted (19-DEC-1999) Production Sequencing Facility, DOE Joint  
JOURNAL Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA  
3 (bases 1 to 186418)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
REFERENCE  
AUTHORS Direct Submission  
TITLE Submitted (01-AUG-2000) DOE Joint Genome Institute, 2800 Mitchell  
JOURNAL Drive, Walnut Creek, CA 94598, USA  
4 (bases 1 to 186418)  
DOE Joint Genome Institute and Stanford Human Genome Center.  
REFERENCE  
AUTHORS Direct Submission  
TITLE Submitted (03-OCT-2001) DOE Joint Genome Institute, 2800 Mitchell  
JOURNAL Drive, Walnut Creek, CA 94598, USA  
On Oct 3, 2001 this sequence version replaced gi:9625331.  
Draft Sequence Produced by DOE Joint Genome Institute  
www.jgi.doe.gov  
Finishing Completed at Stanford Human Genome Center  
www-shgc.stanford.edu  
Quality: Phrap Quality >=40 99.4% of Sequence;  
Estimated Total Number of Errors is 1.1.

FEATURES  
source  
1.186418  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="5"  
/clone="CTB-4518"

ORIGIN  
Query Match 1.7%; Score 52; DB 8; Length 186418;  
Best Local Similarity 100.0%; Pred. No. 5.7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GATCACCTGAGGCGAGAGTTGAGAGCCAGCTCGGCCAACATAGCAAAACC 2950  
|||||



gap	/note="assembly_fragment" 4410. .4509 /estimated_length=100
misc_feature	4510. .6019 /note="assembly_fragment" 6020. .6119 /estimated_length=100
misc_feature	6120. .8353 /note="assembly_fragment" 8354. .8453 /estimated_length=100
misc_feature	8454. .11512 /note="assembly_fragment" 11513. .11612 /estimated_length=100
misc_feature	11613. .13561 /note="assembly_fragment" 13562. .14061 /estimated_length=100
gap	14062. .15526 /note="assembly_fragment" 15527. .15626 /estimated_length=100
misc_feature	15627. .18439 /note="assembly_fragment" 18440. .18539 /estimated_length=100
gap	18540. .20286 /note="assembly_fragment" clone_end:SP6 vector_side:right"
gap	20287. .20386 /estimated_length=100
misc_feature	20387. .24011 /note="assembly_fragment" 24012. .24111 /estimated_length=100
gap	24112. .27768 /note="assembly_fragment" 27769. .27868 /estimated_length=100
misc_feature	27869. .31346 /note="assembly_fragment" 31347. .31446 /estimated_length=100
gap	31447. .35421 /note="assembly_fragment" clone_end:right"
gap	35422. .35521 /estimated_length=100
misc_feature	35522. .40318 /note="assembly_fragment" 40319. .40418 /estimated_length=100
misc_feature	40419. .47211 /note="assembly_fragment" 47212. .47311 /estimated_length=100
gap	

	Query Match	Best Local Similarity	1.7%;	Score 52;	DB 14;	Length 186448;
	Matches	52;	Conservative	0;	Mismatches	0;
					Indels	0;
					Gaps	0;
OY	3071	CAAGATTGTCACCTGCACTTCACGCTTGCGCAACAGACAGAAACTCTCTCTTC				3122
db	148872	CAAGATTGTCACCTGCACTTCACGCTTGCGCAACAGACAGAAACTCTCTCTTC				148923

RESULT	168
AC120045	
LOCUS	188586 bp DNA linear PRI 11-FEB-2003
DEFINITION	Homo sapiens chromosome 15, clone RP11-93209, complete sequence.

ACCESSION	AC120045	GI:28301998
VERSION	AC120045.19	
KEYWORDS	HTG.	
SOURCE	Homo sapiens	(human)
ORGANISM	Homo sapiens	
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.	
AUTHORS	1. (bases 1 to 188596)	
TITLE	Birten,B., Nusbaum,C. and Lander,E.	
JOURNAL	Homo sapiens chromosome 15, clone RP11-93209	
REFERENCE	Unpublished	
AUTHORS	2. (bases 1 to 188596)	
	Birten,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N.,	

**TITLE** Direct Submission  
**JOURNAL** Submitted (02-MAY-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
**REFERENCE** 3 (bases 1 to 188596)  
**AUTHORS** Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,

**TITLE** Direct Submission  
**JOURNAL** Submitted (14-JAN-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
**REFERENCE** 4 (bases 1 to 188596)  
**AUTHORS** Birren, B., Nusbaum, C., Lander, E., Aboueleil, A., Allen, N., Anderson, S., Arachchi, H.M., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bonkagaiter, B., Camarata, J., Chang, J., Choepl, Y., Collymore, A., Cook, A., Cooke, P., Corum, B., Deatellano, K., Diaz, J.S., Dodge, S., Dooley, K., Dorris, L., Erickson, J., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardna, S., Graham, L., Grand-Pierre, N., Half, N., Haploian, D., Hages, B., Hall, J., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kells, C., Lander, T., Levine, R., Lindblad-Toh, K., Liu, G., Lui, A., Mabbitt, R., Maclean, C., Macdonald, P., Major, J., Manning, J., Matthews, C., McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J.,

TITLE  
JOURNAL  
COMMENT

Nguyen, C., Nicol, R., Norbu, C., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Rachupka, A., Ramaamy, U., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, U., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Stubbs, M., Talamas, J., Testaye, S., Theodore, D., Toplam, K., Travers, M., Vassiliev, H., Venkataraman, V. S., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (11-FEB-2003) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Feb 11, 2003 this sequence version replaced gi:27733957.  
All repeats were identified using RepeatMasker:  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: MIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
Project Information  
Center project name: L26756  
Center clone name: 932\_O\_9

# FEATURES

source  
1..188596  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="15"  
/map="15"  
/clone="RP11-93209"  
/clone\_11b="RPCT-11 Human Male BAC"  
repeat\_region  
complement(1..362)  
/rpt\_family="L1PAL5-16"  
repeat\_region  
complement(363..1097)  
/rpt\_family="L1PAB8"  
repeat\_region  
complement(1152..1616)  
/rpt\_family="L1PAB8"  
repeat\_region  
complement(1616..2183)  
/rpt\_family="L1PAB8"  
repeat\_region  
2198..4496  
/rpt\_family="L1PAB8"  
repeat\_region  
complement(4499..5744)  
/rpt\_family="L1PAB8"  
repeat\_region  
5899..5921  
/rpt\_family="AT-rich"  
repeat\_region  
5993..6243  
/rpt\_family="L1MC/D"  
repeat\_region  
6322..6493  
/rpt\_family="MIR"  
repeat\_region  
complement(6832..7032)  
/rpt\_family="MER20"  
repeat\_region  
7175..7212  
/rpt\_family="(TG)n"  
repeat\_region  
7365..7645  
/rpt\_family="AluSg"  
repeat\_region  
complement(7646..7746)  
/rpt\_family="L1MC4a"  
repeat\_region  
8077..8255  
/rpt\_family="MLT1L"  
repeat\_region  
complement(8615..8709)  
/rpt\_family="MIR"  
repeat\_region  
complement(8718..9169)  
/rpt\_family="LOR1b"  
repeat\_region  
10248..10570  
/rpt\_family="AluSg"  
repeat\_region  
complement(11192..11264)  
/rpt\_family="MIR"  
repeat\_region  
complement(11290..11582)  
/rpt\_family="Aluub"  
repeat\_region  
complement(11606..12358)  
/rpt\_family="L1MA3"

repeat\_region complement(12612..14605)  
/rpt\_family="L1MB2"  
repeat\_region complement(14608..14911)  
/rpt\_family="AluUo"  
repeat\_region 15671..15748  
/rpt\_family="MER116"  
repeat\_region complement(16155..16281)  
/rpt\_family="FLAM\_C"  
repeat\_region 16885..17005  
/rpt\_family="AluSc"  
repeat\_region 17157..17208  
/rpt\_family="MER3"  
repeat\_region complement(17168..17217)  
/rpt\_family="MER3"  
repeat\_region 17218..17496  
/rpt\_family="AluUo"  
repeat\_region complement(17497..17654)  
/rpt\_family="MER3"  
repeat\_region complement(17766..17959)  
/rpt\_family="AluUo"  
repeat\_region 17961..18226  
/rpt\_family="MSTa"  
repeat\_region 19462..19581  
/rpt\_family="MIR"  
repeat\_region 20225..20266  
/rpt\_family="MER2"  
repeat\_region complement(20260..20516)  
/rpt\_family="MLT1F2"  
repeat\_region complement(20781..20937)  
/rpt\_family="MLT1F2"  
repeat\_region 20933..21075  
/rpt\_family="MLT1F2"  
repeat\_region 22083..22186  
/rpt\_family="CT-rich"  
repeat\_region complement(22188..22486)  
/rpt\_family="AluSx"  
repeat\_region 22503..22754  
/rpt\_family="L1MB2"  
repeat\_region 22755..23060  
/rpt\_family="AluSx"  
repeat\_region 23063..23297  
/rpt\_family="AluUv"  
repeat\_region 23298..23351  
/rpt\_family="(TA)n"  
repeat\_region 23352..23416  
/rpt\_family="L1MB2"  
repeat\_region 23449..23580  
/rpt\_family="L1MC5"  
repeat\_region complement(23690..23807)  
/rpt\_family="FLAM\_A"  
repeat\_region complement(23834..24139)  
/rpt\_family="AluUv"

Query Match 1.7%; Score 52; DB 8; Length 188596;  
Best Local Similarity 100.0%; Pred.No. 5.7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGAGGTGATCCTGAGGCCGAGGTTGAGACCAAGCTTGCCAACT 2940  
DB 92058 GAGGAGGTGATCCTGAGGCCGAGGTTGAGACCAAGCTTGCCAACT 92109

RESULT 169  
AC146153 188640 bp DNA linear HTG 01-AUG-2003  
LOCUS Pan tciroglydyes chromosome UNK clone RP43-37C16, \*\*\* SEQUENCING IN  
DEFINITION PROGRESS \*\*\*  
ACCESSION AC146153  
VERSION AC146153.1 GI:33387107  
KEYWORDS HTG; HTGS; PHASRI;  
SOURCE Pan tciroglydyes (chimpanzee)  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Pan.  
1 (bases 1 to 188640)  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Wilson, R.K.  
The sequence of Pan troglodytes clone  
Unpublished  
2 (bases 1 to 188640)  
Wilson, R.K.  
Direct Submission  
Submitted (01-AUG-2003) Genetics, Genome Sequencing Center, 4444  
Forest Park Parkway, St. Louis, MO 63108, USA

----- Genome Center -----  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: http://genome.wustl.edu  
Contact: submissions@wustl.edu  
----- Project Information -----  
Center project name: C\_PT037C16

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 18 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1  
1483: contig of 1483 bp in length  
1484  
1583: gap of unknown length  
1584  
3816: contig of 2233 bp in length  
3817  
3916: gap of unknown length  
3917  
6890: contig of 2974 bp in length  
6891  
6990: gap of unknown length  
6991  
9556: contig of 2566 bp in length  
9557  
9657: gap of unknown length  
9657  
12497: contig of 2841 bp in length  
12498  
12597: gap of unknown length  
12598  
15693: contig of 3086 bp in length  
15694  
15793: gap of unknown length  
15794  
18490: contig of 2697 bp in length  
18491  
18590: gap of unknown length  
18591  
24956: contig of 6366 bp in length  
24957  
25056: gap of unknown length  
25057  
28576: contig of 3520 bp in length  
28577  
28676: gap of unknown length  
28677  
34097: contig of 5421 bp in length  
34098  
34197: gap of unknown length  
34198  
41105: contig of 6908 bp in length  
41106  
41205: gap of unknown length  
41206  
53243: contig of 12038 bp in length  
53244  
53343: gap of unknown length  
53344  
64385: contig of 11042 bp in length  
64386  
64485: gap of unknown length  
64486  
77536: contig of 13051 bp in length  
77537  
77636: gap of unknown length  
77637  
92854: contig of 15218 bp in length  
92855  
92954: gap of unknown length  
92955  
185570: contig of 92616 bp in length  
185571  
185670: gap of unknown length  
185671  
187117: contig of 1447 bp in length  
187118  
187217: gap of unknown length  
187218  
188640: contig of 1423 bp in length.

Location/Qualifiers  
1.188640  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/chromosome="UNK"  
/clone="RP43-37C16"  
1.1483  
/note="assembly\_name:Contig13"  
1484.1583

FEATURES  
source  
misc\_feature  
gap

misc\_feature  
1584..3816  
/note="assembly\_name:Contig16"  
gap  
3817..3916  
/estimated\_length=unknown  
misc\_feature  
3917..6890  
/note="assembly\_name:Contig17"  
gap  
6891..6990  
/estimated\_length=unknown  
misc\_feature  
6991..9556  
/note="assembly\_name:Contig18"  
gap  
9557..9656  
/estimated\_length=unknown  
misc\_feature  
9657..12497  
/note="assembly\_name:Contig19"  
gap  
12498..12597  
/estimated\_length=unknown  
misc\_feature  
12598..15693  
/note="assembly\_name:Contig20"  
gap  
15694..15793  
/estimated\_length=unknown  
misc\_feature  
15794..18490  
/note="assembly\_name:Contig21"  
gap  
18491..18590  
/estimated\_length=unknown  
misc\_feature  
18591..24956  
/note="assembly\_name:Contig22"  
gap  
24957..25056  
/estimated\_length=unknown  
misc\_feature  
25057..28576  
/note="assembly\_name:Contig23"  
gap  
28577..28676  
/estimated\_length=unknown  
misc\_feature  
28677..34097  
/note="assembly\_name:Contig24"  
gap  
34098..34197  
/estimated\_length=unknown  
misc\_feature  
34198..41105  
/note="assembly\_name:Contig25"  
gap  
41106..41205  
/estimated\_length=unknown  
misc\_feature  
41206..53243  
/note="assembly\_name:Contig26"  
gap  
53244..53343  
/estimated\_length=unknown  
misc\_feature  
53344..64385  
/note="assembly\_name:Contig27"  
gap  
64386..64485  
/estimated\_length=unknown  
misc\_feature  
64486..77536  
/note="assembly\_name:Contig28"  
gap  
77537..77636  
/estimated\_length=unknown  
misc\_feature  
77637..92854  
/note="assembly\_name:Contig29"  
gap  
92855..92954  
/estimated\_length=unknown  
misc\_feature  
92955..185570  
/note="assembly\_name:Contig30"  
gap  
185571..185670  
/estimated\_length=unknown  
misc\_feature  
185671..187117  
/note="assembly\_name:Contig6"  
gap  
187118..187217  
/estimated\_length=unknown  
misc\_feature  
187218..188640  
/note="assembly\_name:Contig9"

ORIGIN  
Query Match 1.7%; Score 52; DB 14; Length 188640;  
Best Local Similarity 100.0%; Pred. No. 5,7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

0y 2889 GAGGACAGTGCATCTCTGAGGCCAGAGTTGGAACCCAGCTGCGCAACAT 2940  
Db 172431 GAGGACAGTGCATCTCTGAGGCCAGAGTTGGAACCCAGCTGCGCAACAT 172482

RESULT 170  
AC091561

LOCUS AC091561 191037 bp DNA linear PRI 12-JUL-2002  
DEFINITION Homo sapiens chromosome 8, clone RP11-388G22, complete sequence.  
AC091561  
VERSION AC091561.4 GI:16041368  
KEYWORDS HTG;  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 191037)  
Birtren, B., Nusbaum, C. and Lander, E.  
Homo sapiens chromosome 8, clone RP11-388G22  
Unpublished  
2 (bases 1 to 191037)  
Birtren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,  
Barra, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A.,  
Camata, J., Campopiano, A., Chang, J., Choepel, Y., Colangelo, M.,  
Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K.,  
Diaz, J., Dodge, S., Fato, S., Ferreira, P., Fitzgibbon, W., Gage, D.,  
Galagan, J., Gardyna, S., Glende, S., Goyette, M., Graham, L.,  
Grand-Pierre, N., Hagos, B., Heatford, A., Horton, L., Hulme, W.,  
Iliev, I., Johnson, R., Jones, C., Karatas, A., Labocque, K.,  
Lamazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G.,  
Maclean, C., Macdonald, P., Margulis, N., Matthews, C., McCarthy, M.,  
McEwan, P., McKernan, K., McPheters, R., Meldrum, J., Menus, L.,  
Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C.,  
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,  
Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C.,  
Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J.,  
Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupbach, R., Seaman, S.,  
Severy, P., Sougnaz, C., Spencer, B., Strange-Thomann, N.,  
Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Testaye, S.,  
Theodore, J., Travers, M., Travis, N., Triggilo, J., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G.,  
Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission  
JOURNAL Submitted (07-MAY-2001) Whitehead Institute/MIT Center for Genome  
REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA  
AUTHORS 3 (bases 1 to 191037)  
Birtren, B., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,  
Barra, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B.,  
Camata, J., Chang, J., Chararo, B., Choepel, Y., Collymore, A.,  
Cooke, P., Cooke, P., Dearellano, K., Dewar, K., Diaz, J., Dodge, S.,  
Fato, S., Ferreira, P., Fitzgibbon, W., Gage, D., Galagan, J.,  
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,  
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A.,  
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,  
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,  
McCarthy, M., Meldrum, J., Menus, L., Mihova, T., Mlenga, V.,  
Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C.H.,  
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P.,  
Smith, C., Spencer, B., Strange-Thomann, N., Stojanovic, N., Talamas, J.,  
Testaye, S., Theodore, J., Tophan, K., Travers, M., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,  
Zembek, L., Zimmer, A. and Zody, M.

TITLE Direct Submission  
JOURNAL Submitted (12-JUN-2002) Whitehead Institute/MIT Center for Genome  
REFERENCE Research, 320 Charles Street, Cambridge, MA 02141, USA  
COMMENT On Oct 11, 2001 this sequence version replaced gi:15706084.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIR  
Web site: http://www-seg.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L11707  
Center clone name: 388\_G\_22  
-----  
FEATURES  
Source  
Location/Qualifiers  
1. 191037  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"  
/clone="RP11-388G22"  
/clone\_id="RP11-388G22"  
1. 525  
/rpt\_family="L1PA8"  
526. 768  
/rpt\_family="L1PA8"  
1436. 1590  
/rpt\_family="MLT1L"  
complement(3292..3332)  
/rpt\_family="MSTB"  
3333. 3634  
/rpt\_family="ALUSX"  
complement(3635..4004)  
/rpt\_family="MSTB"  
4057. 4201  
/rpt\_family="MIR"  
5412. 5507  
/rpt\_family="MIR"  
6261. 6392  
/rpt\_family="MIR"  
6615. 6698  
/rpt\_family="MER5A"  
8232. 8255  
/rpt\_family="L2"  
8276. 8323  
/rpt\_family="AT\_rich"  
complement(8427..8585)  
/rpt\_family="L2"  
complement(8679..9169)  
/rpt\_family="L1ME2"  
complement(9224..9495)  
/rpt\_family="L1ME"  
complement(9539..10754)  
/rpt\_family="L1ME"  
complement(10892..11295)  
/rpt\_family="L1ME"  
11324. 11633  
/rpt\_family="ALUV"  
complement(12128..12551)  
/rpt\_family="MLT1D"  
13172. 13199  
/rpt\_family="AT\_rich"  
complement(13411..13742)  
/rpt\_family="L2"  
complement(13801..14413)  
/rpt\_family="L2"  
15622. 15941  
/rpt\_family="MER34B"  
complement(15942..16244)  
/rpt\_family="ALUdO"  
16245. 16465  
/rpt\_family="MER34B"  
17295. 17445  
/rpt\_family="MIR"  
17446. 17473  
/rpt\_family="AT\_rich"  
17565. 17587  
/rpt\_family="AT\_rich"



```

repeat_region complement(17591..17704)
repeat_region /rpt_family="PLAM_C"
repeat_region complement(18659..19027)
repeat_region /rpt_family="LRR16A"
repeat_region 21261..21373
repeat_region /rpt_family="GA-rich"
repeat_region 21496..21526
repeat_region /rpt_family="(TG)n"
repeat_region complement(21835..22343)
repeat_region /rpt_family="LIMed"
repeat_region 22937..22971
repeat_region /rpt_family="AT-rich"
repeat_region 23094..23121
repeat_region /rpt_family="AT-rich"
repeat_region complement(23753..23955)
repeat_region /rpt_family="MIR"
repeat_region complement(24750..25038)
repeat_region /rpt_family="AluY"
repeat_region complement(25410..25582)
repeat_region /rpt_family="MIR"
repeat_region complement(25676..26047)
repeat_region /rpt_family="Char11e2a"
repeat_region complement(26056..26269)
repeat_region /rpt_family="L1MA3"
repeat_region complement(26271..29438)
repeat_region /rpt_family="L1MA3"
repeat_region complement(29439..29747)
repeat_region /rpt_family="AluY"
repeat_region complement(29748..31075)
repeat_region /rpt_family="L1MA3"
repeat_region 31076..31098
repeat_region /rpt_family="(TG)n"
repeat_region complement(31099..32198)
repeat_region /rpt_family="L1MA3"
repeat_region 32199..32218
repeat_region /rpt_family="(CAAAA)n"
repeat_region complement(32219..33960)
repeat_region /rpt_family="L1MA3"
repeat_region 34076..34102
repeat_region /rpt_family="(TTA)n"
repeat_region complement(34160..34615)
repeat_region /rpt_family="Char11e2a"
repeat_region complement(34622..34975)
repeat_region /rpt_family="MLT1A2"
repeat_region 34984..35006
repeat_region /rpt_family="AT-rich"
repeat_region complement(35080..35165)
repeat_region /rpt_family="Char11e2a"
repeat_region complement(35297..35528)
repeat_region /rpt_family="Char11e2b"
repeat_region complement(35577..35856)
repeat_region /rpt_family="Char11e2b"
repeat_region complement(36738..36785)
repeat_region /rpt_family="L2"
repeat_region complement(36805..36962)
repeat_region /rpt_family="L3"
repeat_region complement(36953..37063)
repeat_region /rpt_family="L3"
repeat_region 37311..37599
repeat_region /rpt_family="L1MC2"
repeat_region 37635..38296

```

Query Match 1.7%; Score 52; DB 8; Length 191037;  
 Best Local Similarity 100.0%; Fred.No. 5.7e-16;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2899 GATCACTGAGGCGCAGAGTTCAGACGCGCTGGCCCAATAGCGAAACCC 2950  
 |||  
 DB 3391 GATCACTGAGGCGCAGAGTTCAGACGCGCTGGCCCAATAGCGAAACCC 3442

LOCUS AC018681 191234 bp DNA linear HTG 07-JUL-2000  
 DEFINITION Homo sapiens chromosome 8 clone RP11-388622, WORKING DRAFT  
 ACCESSION AC018681.6 GI:8954341  
 VERSION AC018681.6  
 KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homiidae; Homo.  
 1 (bases 1 to 191234)  
 Waterston,R.H.  
 The sequence of Homo sapiens clone  
 Unpublished  
 2 (bases 1 to 191234)  
 Waterston,R.H.  
 Direct Submision  
 Submitted (16-DEC-1999) Genome Sequencing Center, Washington  
 University School of Medicine, 4444 Forest Park Parkway, St. Louis,  
 MO 63108, USA  
 On Jul 7, 2000 this sequence version replaced gi:7231039.

COMMENT

----- Genome Center -----  
 Center: Washington University Genome Sequencing Center  
 Center code: WUGSC  
 Web site: <http://genome.wustl.edu/gsc/index.shtml>  
 ----- Project Information -----  
 Center project name: H\_NH038622  
 ----- Summary Statistics -----  
 Sequencing vector: plasmid; 29k  
 Sequencing vector: M13; 71k  
 Chemistry: Dye-primer ER; 71% of reads  
 Chemistry: Dye-terminator Big Dye; 29% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 186413 bases at least Q40  
 Consensus quality: 188204 bases at least Q30  
 Consensus quality: 189106 bases at least Q20  
 Insert size: 172000; agarose-fp  
 Insert size: 190534; sum-of-contigs  
 Quality coverage: 5.22 in Q20 bases; sum-of-contigs  
 Quality coverage: 5.15 in Q20 bases; sum-of-contigs  
 -----  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 8 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 1724: contig of 1724 bp in length  
 \* 1725 1824: gap of unknown length  
 \* 1825 3510: contig of 1686 bp in length  
 \* 3511 3610: gap of unknown length  
 \* 3611 11673: contig of 8063 bp in length  
 \* 11674 11773: gap of unknown length  
 \* 11774 23956: contig of 12183 bp in length  
 \* 23957 24056: gap of unknown length  
 \* 24057 40166: contig of 16110 bp in length  
 \* 40167 40266: gap of unknown length  
 \* 40267 77292: contig of 37026 bp in length  
 \* 77293 125869: gap of unknown length  
 \* 125870 125969: contig of 48477 bp in length  
 \* 125970 191234: contig of 65265 bp in length.  
 Location/Qualifiers  
 1..191234  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="8"  
 /clone="RP11-388622"

REFERENCE  
AUTHORS 1 (bases 1 to 191540)  
TITLE Ali,J., Bielicki,L. and Cotton,M.  
JOURNAL The sequence of Pan troglodytes BAC clone RP43-2J12  
REFERENCE 2 (bases 1 to 191540)  
AUTHORS Unpublished (2001)  
TITLE 3 (bases 1 to 191540)  
JOURNAL Suleston,J.E. and Wilson,R.  
REFERENCE Unpublished (2001)  
AUTHORS 3 (bases 1 to 191540)  
JOURNAL Wilson,R.K.  
TITLE Direct Submission  
JOURNAL Submitted (17-MAY-2003) Genetics, Genome Sequencing Center, 4444  
Forest Park Parkway, St. Louis, MO 63108, USA  
REFERENCE 4 (bases 1 to 191540)  
AUTHORS Wilson,R.  
TITLE Direct Submission

RESULT	173
LOCUS	AC068676
DEFINITION	AC068676 191866 bp DNA linear HTG 16-DEC-2000 Homo sapiens clone RP11-768N8, WORKING DRAFT SEQUENCE, 14 unordered pieces.
ACCESSION	AC068676
VERSION	AC068676.3 GI:11875293
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

```

REFERENCE 1 (bases 1 to 191866)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens, clone RP11-768N8
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 191866)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Boduriansky,L., Boulanger,B., Brown,A., Burkett,G.,
Boudreau,S., Baldwin,J., Barna,N., Bastien,V., Bede,F.,
Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G.,
Campiano,A., Castle,A., Choepel,Y., Colangelo,M., Collins,S.,
Collamore,A., Cooke,P., Dekrelano,K., Dewar,K., Diaz,J.S.,
Dodge,S., Domingo,M., Doyle,M., Ferreira,P., Fitzhugh,W., Gage,D.,
Galagan,J., Gardyna,S., Glinde,S., Goyette,M., Graham,L.,
Galand-Pierre,N., Grant,G., Hagos,B., Heatford,A., Horton,L.,
Holland,J.C., Iliev,I., Johnson,R., Jones,C., Kam,L., Karacas,A.,
Klein,J., Labocque,K., Lamazares,R., Lander,C., Lehotzky,J.,
Levine,R., Lien,C., Liu,G., Locke,K., MacDonald,P., Margulis,N.,
McCarthy,M., McEwan,P., McGuire,A., McKernan,K., McNeeters,R.,
McDermid,J., Menus,L., Mihova,T., Miranda,C., Mlenga,V., Morrow,J.,
Murphy,T., Naylor,J., Norman,C.H., O'Connor,T., O'Donnell,P.,
O'Neill,D., Oliver,T.M., Oliver,J., Peterson,K., Pierre,N.,
Pisani,C., Pollara,V., Raymond,C., Riley,R., Rogov,P., Rothman,D.,
Roy,A., Santos,R., Schauer,S., Severy,P., Spencer,B.,
Stange-Thomann,N., Stojanovic,N., Subramanian,A., Talamas,J.,
Teefaye,S., Theodore,J., Titrill,A., Travers,M., Triggillo,J.,
Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J.,
Young,G., Zainoun,J., Zimmer,A. and Zody,M.
Direct Submission
Submitted (06-MAY-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Dec 16, 2000 this sequence version replaced gi:10440660.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www-seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu

----- Project Information
Center project name: L10325
Center clone name: 768_N_8
----- Summary Statistics
Sequencing vector: M13; M77815; 2% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 18620 bases at least Q40
Consensus quality: 188219 bases at least Q30
Consensus quality: 189180 bases at least Q20
Insert size: 188000; agarose-fp
Insert size: 190566; sum-of-contigs
Quality coverage: 9.2 in Q20 bases; agarose-fp
Quality coverage: 9.1 in Q20 ba.
NOTE: This is a 'working draft' sequence. It currently
* consists of 14 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1 527: contig of 527 bp in length
* 528 627: gap of 100 bp
* 628 1287: contig of 660 bp in length
* 1288 1387: gap of 100 bp
* 1388 1833: contig of 446 bp in length
* 1834 1933: gap of 100 bp
* 1934 2723: contig of 790 bp in length
* 2724 2823: gap of 100 bp
* 2824 4471: contig of 1648 bp in length
* 4472 4571: gap of 100 bp
* 4572 8184: contig of 3613 bp in length
* 8185 8284: gap of 100 bp

----- FEATURES
source
1. 191866
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="RP11-768N8"
/clone_lib="RP11-11 Human Male BAC"
1. 527
/note="assembly_fragment"
/clone_end:527
vector:hide:left"
528..627
/estimated_length=100
628..1287
/note="assembly_fragment"
1288..1387
/estimated_length=100
1388..1833
/note="assembly_fragment"
1834..1933
/estimated_length=100
1934..2723
/note="assembly_fragment"
2724..2823
/estimated_length=100
2824..4471
/note="assembly_fragment"
4472..4571
/estimated_length=100
4572..8184
/note="assembly_fragment"
8185..8284
/estimated_length=100
8285..18881
/note="assembly_fragment"
18882..18981
/estimated_length=100
18982..33829
/note="assembly_fragment"
33830..33929
/estimated_length=100
33930..51797
/note="assembly_fragment"
51798..51897
/estimated_length=100
51898..67219
/estimated_length=100
67220..67319
/note="assembly_fragment"
67320..67331
/estimated_length=100
67332..88331
/estimated_length=100
88332..132689
/estimated_length=100
132690..132789
/estimated_length=100
132790..170094
/estimated_length=100
170095..191866: contig of 21672 bp in length.
Location/Qualifiers

```

```
gap /note="assembly_fragment"
      170095..170194
      /estimated_length=100
misc_feature /note="assembly_fragment"
      170195..191866
      /clone_end=T7
      vector_side:right"

ORIGIN
Query Match 1.7%; Score 52; DB 14; Length 191866;
Best Local Similarity 100.0%; Pred. No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGAGGTGATCACTGAGGCGAGGAGTTGAGACCACTGGCCAAACAT 2940
Db 170408 GAGGAGGTGATCACTGAGGCGAGGAGTTGAGACCACTGGCCAAACAT 170459

RESULT 174
AC091440/C 191924 bp DNA linear HTG 26-MAY-2001
LOCUS Homo sapiens chromosome UNK clone RP11-561D8, WORKING DRAFT
DEFINITION
AC091440
AC091440.2 GI:14209805
VERSION HTG; HTGS PHASE1; HTGS_DRAFT.
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homidae; Homo.
1 (bases 1 to 191924)
Waterston,R.H.
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 191924)
Waterston,R.H.
Direct Submission
Submitted (20-APR-2001) Genome Sequencing Center, Washington
University School of Medicine, 4444 Forest Park Parkway, St. Louis,
MO 63108, USA
On May 26, 2001 this sequence version replaced gi:13699798.

COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H NH0561D08
----- Summary Statistics -----
Sequencing vector: M13; 7%
Sequencing vector: plasmid; 93%
Chemistry: Dye-primer ET; 7% of reads
Chemistry: Dye-terminator Big Dye; 93% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 177613 bases at least Q40
Consensus quality: 182658 bases at least Q30
Consensus quality: 185159 bases at least Q20
Insert size: 204000; agarose-fp
Insert size: 189524; sum-of-contigs
Quality coverage: 3.86 in Q20 bases; agarose-fp
Quality coverage: 4.21 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 2553: contig of 2553 bp in length
* 2554 2653: gap of unknown length
```

```
* 2654 5221: contig of 2568 bp in length
* 5222 5321: gap of unknown length
* 5322 8504: contig of 3183 bp in length
* 8505 8604: gap of unknown length
* 8605 11131: contig of 2527 bp in length
* 11132 11231: gap of unknown length
* 11232 14456: contig of 3225 bp in length
* 14457 14556: gap of unknown length
* 14557 19951: contig of 5395 bp in length
* 19952 20051: gap of unknown length
* 20052 24638: contig of 4587 bp in length
* 24639 24738: gap of unknown length
* 24739 28963: contig of 4225 bp in length
* 28964 29063: gap of unknown length
* 29064 35750: contig of 6687 bp in length
* 35751 35850: gap of unknown length
* 35851 41905: contig of 6055 bp in length
* 41906 42005: gap of unknown length
* 42006 48503: contig of 6498 bp in length
* 48504 48603: gap of unknown length
* 48604 56276: contig of 7673 bp in length
* 56277 56376: gap of unknown length
* 56377 63617: contig of 7241 bp in length
* 63618 63717: gap of unknown length
* 63718 70545: contig of 6828 bp in length
* 70546 70645: gap of unknown length
* 70646 79240: contig of 8595 bp in length
* 79241 79340: gap of unknown length
* 79341 87568: contig of 8228 bp in length
* 87569 98108: contig of 10440 bp in length
* 98109 98208: gap of unknown length
* 98209 111884: contig of 13676 bp in length
* 111885 111984: gap of unknown length
* 111985 125202: contig of 13218 bp in length
* 125203 125302: gap of unknown length
* 125303 139501: contig of 14199 bp in length
* 139502 139601: gap of unknown length
* 139602 159852: contig of 20251 bp in length
* 159853 159952: gap of unknown length
* 159953 186423: contig of 26371 bp in length
* 186424 186423: gap of unknown length
* 186424 188357: contig of 1934 bp in length
* 188358 188457: gap of unknown length
* 188458 190113: contig of 1656 bp in length
* 190114 190213: gap of unknown length
* 190214 191924: contig of 1711 bp in length.
FEATURES
source
1..191924
location/Qualifiers
organism="Homo sapiens"
mol_type="genomic DNA"
db_xref="taxon:9606"
chromosome="UNK"
clone="RP11-561D8"
1..2553
/note="assembly_name:Contig10"
2554..2653
/estimated_length=unknown
2654..5221
/note="assembly_name:Contig11"
5222..5321
/estimated_length=unknown
5322..8504
/note="assembly_name:Contig12"
8505..8604
/estimated_length=unknown
8605..11131
/note="assembly_name:Contig13"
11132..11231
/estimated_length=unknown
11232..14456
/note="assembly_name:Contig14"
14457..14556
```

```

misc_feature /estimated_length=unknown
14557..19951
/notes="assembly_name:Contig15"
gap 19952..20051
/estimated_length=unknown
misc_feature 20052..24638
/notes="assembly_name:Contig16"
gap 24639..24738
/estimated_length=unknown
misc_feature 24739..28963
/notes="assembly_name:Contig17"
gap 28964..29063
/estimated_length=unknown
misc_feature 29064..35750
/notes="assembly_name:Contig18"
gap 35751..35850
/estimated_length=unknown
misc_feature 35851..41905
/notes="assembly_name:Contig19"
gap 41906..42005
/estimated_length=unknown
misc_feature 42006..48503
/notes="assembly_name:Contig20"
gap 48504..48603
/estimated_length=unknown
misc_feature 48604..56276
/notes="assembly_name:Contig21"
gap 56277..56376
/estimated_length=unknown
misc_feature 56377..63617
/notes="assembly_name:Contig22"
gap 63618..63717
/estimated_length=unknown
misc_feature 63718..70545
/notes="assembly_name:Contig23"
gap 70546..70645
/estimated_length=unknown
misc_feature 70646..79240
/notes="assembly_name:Contig24"
gap 79241..79340
/estimated_length=unknown
misc_feature 79341..87568
/notes="assembly_name:Contig25"
gap 87569..87668
/estimated_length=unknown
misc_feature 87669..98108
/notes="assembly_name:Contig26"
gap 98109..98208
/estimated_length=unknown
misc_feature 98209..111884
/notes="assembly_name:Contig27"
gap 111885..111984
/estimated_length=unknown
misc_feature 111985..125202
/notes="assembly_name:Contig28"
gap 125203..125302
/estimated_length=unknown
misc_feature 125303..1339501
/notes="assembly_name:Contig29"
gap 139502..139601
/estimated_length=unknown
misc_feature 139602..159852
/notes="assembly_name:Contig30"
gap 159853..159952
/estimated_length=unknown
misc_feature 159953..186323
/notes="assembly_name:Contig31
clone_end:T7
vector_side:left"
gap 186324..186423
/estimated_length=unknown
misc_feature 186424..188357
/notes="assembly_name:Contig7"

```

```

gap 188358..188457

Query Match 1.7% Score 52; DB 14; Length 191924;
Best Local Similarity 100.0%; Pred.No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCACATGCACCTCCAGCCTGGGCAAGAGCAAGACTGTCTC 3122
|||||
Db 123776 CAAGTTTGCCACATGCACCTCCAGCCTGGGCAAGAGCAAGACTGTCTC 123725

RESULT 175
AL954212 192328 bp DNA linear PRI 16-OCT-2003
LOCUS Pan troglodytes chromosome 22 clone RP43-043L10 map 22q22.11,
DEFINITION complete sequence.
ACCESSION AL954212
VERSION AL954212.1 GI:37605795
KEYWORDS HTG.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Pan.
1 (bases 1 to 192328)

REFERENCE
AUTHORS The Chimpanzee Chromosome 22 Sequencing Consortium
CONTRM DNA sequence of chimpanzee chromosome 22 and its evolutionary
TITLE Implications
JOURNAL 2 (bases 1 to 192328)
UNPUBLISHED
AUTHORS Kube,M., Sudbrak,R., Borzym,K., Lehrack,S., Thiel,J., Mueller,I.,
Klages,S., Gimel,V., Beck,A., Ben Kahla,A., Lehrach,H., Yaspo,M.L.
and Reinhardt,K.
Direct Submissions
Submitted (18-DEC-2002) Max-Planck-Institute for Molecular
Genetics, Department of Vertebrate Genomics, Ihnestrasse 73,
Berlin, 14195 Germany
The Chimpanzee Chromosome 22 Sequencing Consortium consists of:
*Chinese National Human Genome Center at Shanghai,
Shanghai, China;
*GBF, Dept. of Genome Analysis, Braunschweig, Germany;
*Institute of Molecular Biotechnology, Jena, Germany;
*KRIBB Genome Research Center, Daejeon, Korea;
*Max-Planck-Institute for Molecular Genetics, Berlin, Germany;
*National Institute of Genetics, Mishima, Japan;
*National Yang Ming University Genome Research Center, Taipei,
Taiwan;
*RIKEN Genomic Sciences Center, Yokohama, Japan.

----- Genome Center
Center: Max-Planck-Institute for Molecular Genetics
Center code: MPIMG
----- Project Information
Center clone name: RP43-043L10
----- Summary Statistics
Sequencing vector: pUC18; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 192151 bases at least Q40
Consensus quality: 192300 bases at least Q30
Consensus quality: 192327 bases at least Q20
Quality coverage: 10.18

-----
This sequence was finished as follows unless otherwise noted: all
regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest. -----
Neighboring sequence information:
This clone is overlapped by
PTB-034G05

```



```

/standard_name="OTTHUMP0000023448"
/notes="match: proteins: Q9HAA5"
/codon_start=1
/product="ectodysplasin A2 receptor"
/protein_id="CAH70433.1"
/db_xref="GI:5563359"
/db_xref="GOA:Q5VYX0"
/db_xref="InterPro:IPR001368"
/translation="MDQENYWDQRCVTCORCGPGQGLSKDQCYGEGGAYCTAC
PRRYKSSWGHHCOSCTICAVINRQKNCRTISNAVGDCIPRYRTRIGLQDQ
ECIPCTKQPTSEVQCAFQSLVADPTVPQEQATLVAVSLVFLATLGEFL
YKQFNRHCQSGGLQFADTKAKESLFPVPSKETSASVSNITOTPLNFIIL
EDDCSTSGFPTQSEFTMASCTSESHSHWVHSPIECTEIDLQKFSASATGAEITLG
NVTSETRDLINVPFVPS"
complement(join(103737..104119,104292..104348,
106886..107050,108674..108759,109301..109479,
120187..120273))
/genes="EDA2R"
/locus_tag="RP11-133K18.2-002"
/standard_name="OTTHUMP0000023449"
/notes="match: proteins: Q81ZA6"
/codon_start=1
/product="ectodysplasin A2 receptor"
/protein_id="CAH70434.1"
/db_xref="GI:5563360"
/db_xref="GOA:Q5VYX9"
/db_xref="InterPro:IPR001368"
/translation="MDQENYWDQRCVTCORCGPGQGLSKDQCYGEGGAYCTAC
PRRYKSSWGHHCOSCTICAVINRQKNCRTISNAVGDCIPRYRTRIGLQDQ
ECIPCTKQPTSEVQCAFQSLVADPTVPQEQATLVAVSLVFLATLGEFL
YKQFNRHCQSGGLQFADTKAKESLFPVPSKETSASVSNITOTPLNFIIL
EDDCSTSGFPTQSEFTMASCTSESHSHWVHSPIECTEIDLQKFSASATGAEITLG
NVTSETRDLINVPFVPS"

```

## ORIGIN

Query Match 1.7% Score 52; DB 8; Length 192505;  
 Best Local Similarity 100.0%; Pred. No. 5,7e-16;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGCGACCTGCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
 |||||  
 DB 128720 CAAGATTGCGACCTGCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 128771

## RESULT 177

## AC110601/c

DEFINITION Homo sapiens chromosome 15 clone RP11-422F16 map 15, 4 unordered  
 pieces

AC110601 193351 bp DNA linear HTG 27-MAR-2003  
 AC110601.4 GI:25103366  
 VERSION HTG; HTGS\_PHASE1; HTGS\_FULLTOP; HTGS\_CANCELLED.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homiidae; Homo.

REFERENCE 1 (bases 1 to 193351)  
 AUTHORS Birren, B., Nussbaum, C. and Lander, E.  
 TITLE Homo sapiens chromosome 15, clone RP11-422F16  
 JOURNAL Unpublished  
 REFERENCE 2 (bases 1 to 193351)  
 AUTHORS Birren, B., Linton, L., Nussbaum, C., Lander, E., Ali, A., Allen, N.,  
 Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Bouhgalter, B.,  
 Brown, A., Camarata, J., Campiolo, A., Chang, J., Chazaro, B.,  
 Chopel, Y., Collangelo, M., Collins, S., Collymore, A., Cook, A.,  
 Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S.,  
 Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J., Gardyna, S.,  
 Ginde, S., Gord, S., Goyette, M., Graham, M., Grand-Pierre, N.,  
 Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
 Kamat, A., Karatas, A., Kells, C., Lacombe, K., Lamazares, R.,

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 Submitted (13-FEB-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 193351)  
 Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,  
 Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Bouhgalter, B.,  
 Camarata, J., Chang, J., Chazaro, B., Chopel, Y., Collymore, A.,  
 Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,  
 Faro, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,  
 Gardyna, S., Gord, S., Graham, M., Grand-Pierre, N., Hafez, N.,  
 Hago, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
 Kamat, A., Karatas, A., Kells, C., Lander, E., Levine, R.,  
 Lindblad-Toh, K., Liu, G., Maclean, C., MacDonald, P., Major, J.,  
 Matthews, C., McCarthy, M., Meldrum, J., Menuez, L., Mhova, T.,  
 Miengo, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Noru, C.,  
 Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,  
 Peterson, K., Phunhphang, P., Piere, N., Raymond, C., Retta, R.,  
 Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R.,  
 Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N.,  
 Stojanovic, N., Talamas, J., Testafaye, S., Theodore, J., Tophan, K.,  
 Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,  
 Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
 JOURNAL  
 COMMENT  
 Submitted (27-MAR-2003) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Nov 19, 2002 this sequence version replaced gi:22507177.  
 All repeats were identified using RepeatMasker:  
 Smit, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html

Center: Whitehead Institute/ MIT Center for Genome Research  
 Direct Submission  
 Center code: MIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: L25406  
 Center clone name: 422\_F\_16  
 \*\*\*\*\*  
 \* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 4 contigs. The true order of the pieces  
 \* is not known and their order in this sequence record is  
 \* arbitrary. Gaps between the contigs are represented as  
 \* runs of N, but the exact sizes of the gaps are unknown.  
 \* This record will be updated with the finished sequence  
 \* as soon as it is available and the accession number will  
 \* be preserved.  
 \* 1 23397: contig of 23397 bp in length  
 \* 23398 23497: gap of 100 bp  
 \* 23498 85271: contig of 61874 bp in length  
 \* 85172 85271: gap of 100 bp  
 \* 85272 140698: contig of 55427 bp in length  
 \* 140699 140798: gap of 100 bp  
 \* 140799 193351: contig of 52553 bp in length.  
 Location/Qualifiers  
 1..193351  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="15"  
 /map="15"





Matthews, C., McCarthy, M., Meldim, J., Meneus, L., Mihova, T.,  
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,  
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,  
Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,  
Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnupack, R.,  
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thumann, N.,  
Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,  
Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,  
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (01-NOV-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 194189)  
Barron, B., Baertien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,  
Barra, N., Nusbam, C., Lander, E., Ali, A., Allen, N., Anderson, S.,  
Camarata, J., Chang, J., Chazaro, B., Choepel, T., Collamore, A.,  
Cook, A., Cooke, P., Dearlilano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,  
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N.,  
Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C.,  
Kamat, A., Karatae, A., Kelle, C., Landers, T., Levine, R.,  
Limblad-Ton, K., Liu, G., Maclean, C., MacDonald, P., Major, J.,  
Matthews, C., McCarthy, M., Meldim, J., Meneus, L., Mihova, T.,  
Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C.,  
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J.,  
Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R.,  
Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schnupack, R.,  
Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thumann, N.,  
Stojanovic, N., Talamas, J., Tesfaye, S., Theodore, J., Topham, K.,  
Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X.,  
Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.  
Direct Submission  
Submitted (07-JAN-2003) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jan 7, 2003 this sequence version replaced gi:24462416.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence.submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L21594  
Center clone name: 55\_U\_15  
-----  
FEATURES  
source  
-----  
Location/Qualifiers  
1. 194189  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="8"  
/map="8"  
/clone="RP11-55J15"  
/clone\_1lb="RP11-11 Human Male BAC"  
71. 416  
/rpt\_family="AluSx"  
complement(773. 1076)  
repeat\_region  
/rpt\_family="AluUo"  
1725. 1776  
repeat\_region  
/rpt\_family="MIR"  
2452. 2753  
repeat\_region  
/rpt\_family="AluSg"  
3295. 3420  
repeat\_region  
/rpt\_family="AluSg/x"  
3653. 3780  
repeat\_region  
/rpt\_family="MIR"  
4093. 4386  
repeat\_region  
/rpt\_family="AluSx"  
complement(4392. 4667)  
repeat\_region  
/rpt\_family="AluUo"  
5046. 5088  
repeat\_region  
/rpt\_family="TTTA)n"  
complement(5093. 5372)  
repeat\_region  
/rpt\_family="AluSg"  
5425. 5694  
repeat\_region  
/rpt\_family="L2"  
complement(5697. 5767)  
repeat\_region  
/rpt\_family="MIR"  
6202. 6432  
repeat\_region  
/rpt\_family="MIR"  
complement(6511. 6810)  
repeat\_region  
/rpt\_family="AluY"  
7006. 7115  
repeat\_region  
/rpt\_family="MIR"  
7276. 7753  
repeat\_region  
/rpt\_family="MLT1D"  
complement(10040. 10142)  
repeat\_region  
/rpt\_family="MIR"  
complement(10343. 10667)  
repeat\_region  
/rpt\_family="AluSx"  
10734. 10851  
repeat\_region  
/rpt\_family="TTA)n"  
complement(11143. 11453)  
repeat\_region  
/rpt\_family="AluSx"  
11753. 11822  
repeat\_region  
/rpt\_family="MLT1C"  
complement(11823. 12124)  
repeat\_region  
/rpt\_family="AluSx"  
12125. 12516  
repeat\_region  
/rpt\_family="MLT1C"  
12735. 12890  
repeat\_region  
/rpt\_family="MIR"  
13791. 13829  
repeat\_region  
/rpt\_family="L1MA3"  
13830. 14147  
repeat\_region  
/rpt\_family="AluSg"  
14156. 14457  
repeat\_region  
/rpt\_family="AluSx"  
14458. 14616  
repeat\_region  
/rpt\_family="L1MA3"  
15897. 15952  
repeat\_region  
/rpt\_family="L2"  
complement(15963. 16079)  
repeat\_region  
/rpt\_family="L2"  
complement(16310. 16421)  
repeat\_region  
/rpt\_family="L2"  
16348. 16614  
repeat\_region  
/rpt\_family="MIR"  
16927. 16967  
repeat\_region  
/rpt\_family="TG)n"  
17335. 17407  
repeat\_region  
/rpt\_family="TCTCTG)n"  
17691. 17766  
repeat\_region  
/rpt\_family="MIR"  
18015. 18167  
repeat\_region  
/rpt\_family="MER5A"  
18277. 18313  
repeat\_region  
/rpt\_family="TG)n"  
18425. 18468  
repeat\_region  
/rpt\_family="T-rich"  
18542. 18606  
repeat\_region  
/rpt\_family="CT-rich"  
19749. 19961  
repeat\_region  
/rpt\_family="AluUo"  
20354. 20630  
repeat\_region  
/rpt\_family="CT-rich"  
20597. 20634  
repeat\_region  
/rpt\_family="TCTG)n"  
21246. 21269  
repeat\_region  
/rpt\_family="MIR"  
21270. 21595  
repeat\_region  
/rpt\_family="AluSx"  
21596. 21764  
repeat\_region  
/rpt\_family="MIR"

```

repeat_region      21936..22047
repeat_region      /rpf_family="MUR3"
repeat_region      22184..22268
repeat_region      /rpf_family="12"

Query Match
Best Local Similarity 100.0%; Pred. No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3071 CAAGATTGTCACCTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122
Db      134370 CAAGATTGTCACCTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 134421

RESULT 180
AP002026      195389 bp      DNA      linear      PRI 19-MAR-2003
LOCUS      Homo sapiens genomic DNA, chromosome 4q22-q24, clone:429K21,
DEFINITION      complete sequence.
ACCESSION      AP002026
VERSION      AP002026.2 GI:29122862
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.

REFERENCE
1 Tsai, S.
  The Complete Genomic DNA Sequence of the Human ADH Gene Complex 1
  JOURNAL
  TITLE      Unpublished
  AUTHORS      Tsai, S.F.
  DIRECT SUBMISSION
  Submitted (09-MAY-2000) Shih-Feng Tsai, National Yang-Ming
  University, Institute of Genetics, 155 Yi-Rong St. Section 2,
  Peitou, Taipei, Taiwan 11221, Republic of China
  (E-mail:lympeisa@ym.edu.tw, URL:http://genome.ym.edu.tw/,
  Tel:886-2-28267043, Fax:886-2-28264930)
  On Mar 18, 2003 this sequence version replaced gi:9293861.
  Quality: the expected Phred/Phrap calculated error rate (per 10kb)
  is 0.68; Estimated total number of errors is 13.39.
  Location/Qualifiers
            1..195389
            /organism="Homo sapiens"
            /mol_type="Genomic DNA"
            /db_xref="taxon:9606"
            /chromosome="4"
            /map="4q22-q24"
            /clone="429K21"

FEATURES
source

ORIGIN

Query Match
Best Local Similarity 100.0%; Pred. No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3071 CAAGATTGTCACCTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122
Db      85260 CAAGATTGTCACCTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 85311

RESULT 181
AC125238/c      196622 bp      DNA      linear      PRI 30-APR-2005
LOCUS      Homo sapiens BAC clone RP11-711C24 from 2, complete sequence.
ACCESSION      AC125238
VERSION      AC125238.5 GI:22267884
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homnidae; Homo.

```

```

REFERENCE
1 (bases 1 to 196622)
AUTHORS      Nguyen, C., Bielicki, L. and Schatzkamer, K.
TITLE      The sequence of Homo sapiens BAC clone RP11-711C24
JOURNAL      Unpublished (2001)
REFERENCE
2 (bases 1 to 196622)
AUTHORS      Wacreston, R.H.
TITLE      Direct Submission
JOURNAL      Submitted (20-JUN-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE
3 (bases 1 to 196622)
AUTHORS      Wacreston, R.H.
TITLE      Direct Submission
JOURNAL      Submitted (08-AUG-2002) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
REFERENCE
4 (bases 1 to 196622)
AUTHORS      Wacreston, R.
TITLE      Direct Submission
JOURNAL      Submitted (16-AUG-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE
5 (bases 1 to 196622)
AUTHORS      Wacreston, R.
TITLE      Direct Submission
JOURNAL      Submitted (15-OCT-2002) Department of Genetics, Washington
            University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA
REFERENCE
6 (bases 1 to 196622)
AUTHORS      Wilson, R.K.
TITLE      Direct Submission
JOURNAL      Submitted (30-APR-2005) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
On Aug 16, 2002 this sequence version replaced gi:22138709.

----- Genome Center
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu
Contact: submissions@wustl.wustl.edu
----- Summary Statistics
Center project name: H_NH0711C24
-----

NOTICE:
This sequence was finished as follows unless otherwise noted:
all regions were double stranded, sequenced with an alternate
chemistry, or covered by high quality data (i.e., phred quality >=
30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by sequence
from more than one subclone; and the assembly was confirmed by
restriction digest.

MAPPING INFORMATION:
Mapping information for this clone was provided by Dr. Wes Warren,
Department of Genetics, Washington University, St. Louis MO. For
additional information about the map position of this sequence, see
http://genome.wustl.edu

SOURCE INFORMATION:
The RP11-11 human BAC library was made from the blood of one male
donor, as described by Osogawa, K., Woon, P.Y., Zhao, B., Frengen, E.,
Tateno, M., Catanesse, J.J. and de Jong, P.J. (1998) An improved
approach for construction of bacterial artificial chromosome
libraries. Genomics 51:1-8. The clone may be obtained either from
Research Genetics, Inc. (http://www.resgen.com) or Pieter de Jong
and coworkers at http://www.choir1.org
VECTOR: pBACe3.6

NEIGHBORING SEQUENCE INFORMATION:
This sequence is the entire insert of the clone. This clone is
overlapped by AC069314 and AC010138.
Location/Qualifiers
1..196622

```

```

/misc_feature      /organism="Homo sapiens"
/misc_feature      /mol_type="genomic DNA"
/misc_feature      /db_xref="taxon:9606"
/misc_feature      /chromosome="2"
/misc_feature      /clone_1lb="RP11-711C24"
/misc_feature      /clone_1lb="RP11-711C24"
/misc_feature      /note="CpG Island (GC=60.7, o/e=0.67, #CpGs=18)"
/misc_feature      34311..34561
/misc_feature      /note="CpG Island (GC=57.4, o/e=1.17, #CpGs=21)"
/misc_feature      35283..35592
/misc_feature      /note="CpG Island (GC=73.5, o/e=0.70, #CpGs=11)"
/misc_feature      98867..130225
/misc_feature      /gene="CD28"
/misc_feature      join(98867..99140,119025..119381,122040..122164,
/misc_feature      127176..130225)
/misc_feature      /gene="CD28"
/misc_feature      join(99089..99140,119025..119381,122040..122164,
/misc_feature      127176..127304)
/misc_feature      /gene="CD28"
/misc_feature      /note="Homo sapiens CD28 antigen (Tp44) (CD28), mRNA.;
/misc_feature      H.NH0711C24.1
/misc_feature      This gene was based on gi(5453610)"
/misc_feature      /product="unknown"
/misc_feature      /protein_id="AAV24123.1"
/misc_feature      /db_xref="GI:62988736"
/misc_feature      /translation="MDRLALNLPISIVTGKILVKOSPMVAVDNAVNLSCKSY
/misc_feature      NLPSEFRAHLKGLDSAVECVVGNYSQQLQVSKTGNDGKAGNSVFYIOLN
/misc_feature      VYNQDIDYFCKIEVMYPPYLDNENKGTIIHVKGLCPSLFPQSPKPFVILVVG
/misc_feature      GVLACYSLVTVAFIIFWVRSKSRSLHSDVMNMPRRPGPRKHQPYAPPRDFAAY
/misc_feature      RS"
/misc_feature      171163..171427
/misc_feature      /note="CpG Island (GC=62.3, o/e=0.80, #CpGs=20)"
/misc_feature      17943..178329
/misc_feature      /note="CpG Island (GC=73.1, o/e=0.60, #CpGs=34)"

ORIGIN
Query Match      1.7%; Score 52; DB 8; Length 196622;
Best Local Similarity 100.0%; Pred. No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy      3071 CAAGATTGCGACCTGACCTCCAGCTGCGCAACAGAGCAAGCTGTCTC 3122
Db      70272 CAGATTGCGACCTGACCTCCAGCTGCGCAACAGAGCAAGCTGTCTC 70221

RESULT 182
AC099058      196623 bp DNA linear PRI 26-JAN-2002
DEFINITION      Homo sapiens chromosome 3 clone RP11-714G12, complete sequence.
ACCESSION      AC099058
VERSION      AC099058.2 GI:18376904
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 196623)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Phelps, K.A., Raymond, C. and Haugen, E.D.
Direct Submision
Unpublished
2 (bases 1 to 196623)
Kaul, R.K., Olson, M.V., Raymond, C. and Haugen, E.D.
Direct Submision
Submitted (08-NOV-2001) Genome Center, University of Washington,
Box 352145, Seattle, WA 98195, USA
3 (bases 1 to 196623)
Kaul, R.K., Olson, M.V., Zhou, Y., James, R.A., Rouse, G., Wu, Z.,
Saenphimmachak, C., Phelps, K.A., Raymond, C. and Haugen, E.D.
Direct Submision

```

## JOURNAL COMMENT

Submitted (26-JAN-2002) Genome Center, University of Washington,  
Box 352145, Seattle, WA 98195, USA  
On Jan 26, 2002 this sequence version replaced gi:16799016.

## ----- Genome Center

Center: University of Washington Genome Center  
Center Code: UWGC  
Web site: <http://www.genome.washington.edu>  
Contact: [uwgctg@u.washington.edu](mailto:uwgctg@u.washington.edu)  
Drafting Center: BCM

## ----- Project Information

Center project name: chr-3  
Center clone name: RP11-714G12 (bc0534)

## ----- Summary Statistics

Sequencing vector: unknown; 46% of reads  
Chemistry: Dye-terminator BT; 92% of reads  
Chemistry: Dye-terminator Big Dye; 8% of reads  
Assembly program: Phrap; version 0.990319  
Consensus quality: 196528 bases at least Q40  
Consensus quality: 196607 bases at least Q30  
Consensus quality: 196623 bases at least Q20  
Insert size: 196623; sum-of-contigs  
Quality coverage: 10.6x in Q20 bases, sum-of-contigs

## ----- Overlapping Sequences:

5': RP11-757A (UWGC:bc0549) AC104438  
3': RP11-45P4 (UWGC:bc0161) AC015642

## ----- Sequence Quality Assessment:

This entry has been annotated with sequence quality  
estimates computed by the Phrap assembly program.  
All manually edited bases have been reduced to quality zero.  
Quality levels above 40 are expected to have less than  
1 error in 10,000 bp.  
Base-by-base quality values are not generally visible from the  
Genbank flat file format but are available as part  
of this entry's ASN.1 file.

This sequence was finished as follows unless otherwise noted:  
all regions were either double-stranded or sequenced with an  
alternate chemistry or covered by high quality data (i.e., Pired  
quality >= 30); an attempt was made to resolve all sequencing  
problems, such as compressions and repeats; all regions were  
covered by at least one plasmid subclone or more than one M13  
subclone; and the assembly was confirmed by restriction digest.

## ----- Sequence Validation:

This sequence has been validated by Multiple Complete Digest  
fingerprinting. Comparison of the experimentally derived digest  
fragments with sequence-predicted fragments is given below.  
The electronically-digested sequence consists of both insert and  
vector, in order to accurately represent the entire circular BAC.  
Small fragments below a variable cutoff (approximately 400-800 bp)  
are not resolved in the fingerprint and hence do not appear  
in the table. There are no significant remaining discrepancies  
between the experimental and predicted values. Uniquely ordered  
fragments are separated by dashed lines.

## EcoRI

## HindIII

## BglII

SeqDerMap	Fingerprint	SeqDerMap	Fingerprint	SeqDerMap	Fingerprint
8696	8773	873	862	6586	6552
6	<800	6382	6565	2067	2057
1412	1360	512	<800	11841	11792
3175	3219	449	<800	2788	2800
4008	4037	10461	10208	1276	1274

3177	3219	1048	1014	5097	5045
11	<800	48	<800	30195	30333
13	<800	13755	13252	12139	11792
18	<800	1120	1124	2717	2800
152	<800	4425	4426	15035	14805
2523	2608	243	<800	4759	4767
1226	1156	710	<800	352	<800
1584	1667	1273	1209	2393	2384
1748	1667	2375	2359	2558	2580
2184	2196	1606	1605	6014	5959
1165	1156	2656	2848	8265	8530
939	944	14802	14596	1985	1920
187	<800	3065	3205	6151	6165
2562	2608	2314	2428	9322	9390
340	<800	1401	1419	842	857
1077	1069	893	914	3925	3969
1154	1156	2615	2740	772	857
1674	1667	2173	2185	8443	8530
1185	1156	58	<800	3279	3297
26	<800	172	<800	29	<800
3681	3677	998	1014	220	<800
237	<800	1850	1898	10592	10553
4100	4037	4785	4841	1322	1274
8739	8773	126	<800	1407	1376
1356	1360	2675	2848	828	857
1528	1532	1217	1209	1138	1121
2083	2196	1445	1419	5860	5959
6682	6726	4251	4200	153	<800
8269	8217	7428	7543	250	<800
3621	3677	3474	3582	2251	2384
5344	5303	3347	3426	867	857
6076	6046	1463	1419	6660	6731
4071	4037	839	862	4484	4513
776	<800	10565	10208	559	<800
746	794	662	<800	3026	3026
3244	3219	9	<800	10591	10553
6000	6046	2925	3071	34	<800

Query Match 1.7%; Score 52; DB 8; Length 196623;  
 Best Local Similarity 100.0%; Pred. No. 5.7e-16;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2899 GATCACCTGAGGCGGAGAGTTGAGACCGGCTGGCCACATAGCAAAACC 2950  
 Db 127069 GATCACCTGAGGCGGAGAGTTGAGACCGGCTGGCCACATAGCAAAACC 127120

RESULT 183  
 AC026160  
 LOCUS  
 DEFINITION Homo sapiens chromosome 3 clone RP11-1017A6 map 3p, WORKING DRAFT  
 SEQUENCE, 16 unordered pieces.  
 ACCESSION AC026160.2 GI:8121162  
 VERSION  
 KEYWORDS HTG; HTGS; PHASE1; HTGS\_DRAFT.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 1 (bases 1 to 198105)  
 REFERENCE  
 AUTHORS Bao,J., Bao,Q., Bao,W., Bian,X., Cao,T., Chen,C., Chen,J., Ding,H.,  
 Dong,W., Fan,H., Feng,X., Guan,Q., Gu,X., Guo,D., He,L., Hu,S.,  
 Huang,F., Jin,Y., Kang,N., Li,C., Li,G., Li,J., Li,L.,  
 Li,S., Li,T., Liu,Y., Liu,N., Liu,B., Liu,Y., Li,W., Li,W., Li,Y.,  
 Luo,J., Niu,Y., Qi,Q., Qi,X., Song,S., Sun,M., Sun,W., Sun,Y.,  
 Tao,R., Wang,H., Wang,J., Wang,J., Wang,L., Wang,R.,  
 Wang,X., Wang,X., Wang,Y., Wu,D., Wu,Q., Xie,F., Xuan,Z., Xue,Y.,  
 Yan,C., Yang,X., Yu,B., Zeng,Y., Zhang,G., Zhang,H.,  
 Zhang,L., Zhang,M., Zhang,X., Zhang,X., Zhang,Y.,  
 Zhang,Z., Zhu,B., Yu,J. and Yang,H.  
 Chromosome 3p genomic sequence  
 Unpublished  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS Kang,N., Hu,S., Dong,W., Wang,J., Zhang,Y., Zhang,H., Liu,B.,  
 Bao,W., Sun,Y., Wu,Q., Wang,H., Yang,X., Cheng,C., Wang,Y., Niu,Y.,  
 Qi,X., Li,T., Zhang,H., Liu,N., Wu,D., Yu,B., Fan,H., Liu,Y.,  
 Li,G., Li,C., Bao,Q., Bao,J., Wang,X., Song,L., Zhang,L., Guo,D.,  
 Huang,F., Zhang,G., Li,J., Bian,X., Zhang,M., Li,L., Feng,X., Yu,J.  
 and Yang,H.

TITLE  
JOURNAL

Direct Submission  
Submitted (21-MAR-2000) Human Genomic Center, Institute of  
Genetics, Chinese Academy of Sciences, Datun Road, Beijing, Beijing  
100101, P.R.China  
On May 31, 2000 this sequence version replaced gi:7271977.

## COMMENT

-----Genome Center  
Center:Beijing Center  
Center code:Beijing  
Website:http://hgc.igtp.ac.cn  
http://www.genomics.org.cn  
Contact:hgc@igtp.ac.cn

----- Project Information  
Center project name:1# project  
Center clone name: RP11-1017A6

----- Summary Statistics  
Sequencing vector: pUC18; 100% of reads  
Chemistry: Dye-terminator; ET 5% of reads  
Chemistry: Dye-terminator Big Dye; 45% of reads  
Assembly program: Phrap; version 0.990329  
Consensus quality: 189600 bases at least Q40  
Consensus quality: 194382 bases at least Q30  
Consensus quality: 197736 bases at least Q20  
Insert size: 190445; sum-of-Contigs  
Quality coverage: 4.76x in Q20 bases;sum-of-Contigs

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 16 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1194: contig of 1194 bp in length  
\* 1195 1294: gap of unknown length  
\* 1295 3027: contig of 1733 bp in length  
\* 3028 3127: gap of unknown length  
\* 3128 4527: contig of 1400 bp in length  
\* 4528 4627: gap of unknown length  
\* 4628 8413: contig of 3786 bp in length  
\* 8414 8513: gap of unknown length  
\* 8514 13191: contig of 4678 bp in length  
\* 13192 13291: gap of unknown length  
\* 13292 19447: contig of 6156 bp in length  
\* 19448 19547: gap of unknown length  
\* 19548 25715: contig of 6168 bp in length  
\* 25716 25815: gap of unknown length  
\* 25816 36292: contig of 10477 bp in length  
\* 36293 36392: gap of unknown length  
\* 36393 49087: contig of 12695 bp in length  
\* 49088 49187: gap of unknown length  
\* 49188 61854: contig of 12667 bp in length  
\* 61855 61954: gap of unknown length  
\* 61955 75324: contig of 13370 bp in length  
\* 75325 75424: gap of unknown length  
\* 75425 88685: contig of 13661 bp in length  
\* 88686 88785: gap of unknown length  
\* 88786 104410: contig of 15625 bp in length  
\* 104411 104510: gap of unknown length  
\* 104511 120475: contig of 15965 bp in length  
\* 120476 120575: gap of unknown length  
\* 120576 141215: contig of 20639 bp in length  
\* 141215 141315: gap of unknown length  
\* 141315 198105: contig of 56791 bp in length.

## FEATURES

source  
Location/Qualifiers

1..198105  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="3"  
/map="3p"  
/clone="RP11-1017A6"  
misc\_feature  
1..1194

/note="assembly\_name:Contig3"  
1195..1294  
/estimated\_length=unknown  
1295..3027  
/note="assembly\_name:Contig4"  
3028..3127  
/estimated\_length=unknown  
3128..4527  
/note="assembly\_name:Contig5"  
4528..4627  
/estimated\_length=unknown  
4628..8413  
/note="assembly\_name:Contig6"  
8414..8513  
/estimated\_length=unknown  
8514..13191  
/note="assembly\_name:Contig7"  
13192..13291  
/estimated\_length=unknown  
13292..19447  
/note="assembly\_name:Contig8  
clone\_end:SB6  
vector\_side:right"  
19448..19547  
/estimated\_length=unknown  
19548..25715  
/note="assembly\_name:Contig9"  
25716..25815  
/estimated\_length=unknown  
25816..36292  
/note="assembly\_name:Contig10"  
36293..36392  
/estimated\_length=unknown  
36393..49087  
/note="assembly\_name:Contig11"  
49088..49187  
/estimated\_length=unknown  
49188..61854  
/note="assembly\_name:Contig12  
clone\_end:17  
vector\_side:right"  
61855..61954  
/estimated\_length=unknown  
61955..75324  
/note="assembly\_name:Contig13"  
75325..75424  
/estimated\_length=unknown  
75425..88685  
/note="assembly\_name:Contig14"  
88686..88785  
/estimated\_length=unknown  
88786..104410  
/note="assembly\_name:Contig15"  
104411..104510  
/estimated\_length=unknown  
104511..120475  
/note="assembly\_name:Contig16"  
120476..120575  
/estimated\_length=unknown  
120576..141214  
/note="assembly\_name:Contig17"  
141215..141314  
/estimated\_length=unknown  
141315..198105  
/note="assembly\_name:Contig18"

## ORIGIN

Query Match 1.7% Score 52; DB 14; Length 198105;  
Best Local Similarity 100.0%; Pred.No. 5.7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAAGTTGCGACCTGCACTCGAGCTGGGCAAGAGCAAGACTCTGTCTC 3122  
|||||

Db 113417 CAAGATTGTGCACTGCACCTCGAGCTGGGCAAGACAGCAAGACTCTGTCTC 113468

RESULT 184  
AC008108/c  
LOCUS  
DEFINITION AC008108 200000 bp DNA linear HTG 23-JUL-1999  
Homo sapiens chromosome 4, \*\*\* SEQUENCING IN PROGRESS \*\*\*, 17  
unnumbered pieces.

ACCESSION AC008108  
VERSION AC008108.1 GI:5578727  
KEYWORDS HTG; HTGS\_PHASE1.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.

REFERENCE 1 (bases 1 to 200000)  
AUTHORS Stone,N.E., Schmutz,J.J., Cox,D.R. and Myers,R.M.  
TITLE Direct Submission  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 200000)  
AUTHORS Stone,N.E., Schmutz,J.J., Cox,D.R. and Myers,R.M.  
TITLE Direct Submission  
JOURNAL Submitted (23-JUL-1999) Department of Genetics, Stanford Human  
Genome Center, 855 California Avenue, Palo Alto, CA 94304, USA

\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 17 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.  
\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.

1 1205: contig of 1205 bp in length  
\* 1 1206 12031: gap of unknown length  
\* 12032 13178: contig of 1147 bp in length  
\* 13179 24004: gap of unknown length  
\* 24005 25125: contig of 1121 bp in length  
\* 25126 35951: gap of unknown length  
\* 35952 37175: contig of 1224 bp in length  
\* 37176 48001: gap of unknown length  
\* 48002 49547: contig of 1546 bp in length  
\* 49548 60373: gap of unknown length  
\* 60374 61456: contig of 1083 bp in length  
\* 61457 72281: gap of unknown length  
\* 72282 73415: contig of 1134 bp in length  
\* 73416 84240: gap of unknown length  
\* 84241 85493: contig of 1253 bp in length  
\* 85494 96318: gap of unknown length  
\* 96319 97555: contig of 1237 bp in length  
\* 97556 108380: gap of unknown length  
\* 108381 110360: contig of 1980 bp in length  
\* 110361 121185: gap of unknown length  
\* 121186 122779: contig of 1593 bp in length  
\* 122780 133603: gap of unknown length  
\* 133604 135588: contig of 1985 bp in length  
\* 135589 146413: gap of unknown length  
\* 146414 148101: contig of 1688 bp in length  
\* 148102 158926: gap of unknown length  
\* 158927 160582: contig of 1656 bp in length  
\* 160583 171407: gap of unknown length  
\* 171408 173903: contig of 2496 bp in length  
\* 173904 184728: gap of unknown length  
\* 184729 186679: contig of 1951 bp in length  
\* 186680 197504: gap of unknown length  
\* 197505 200000: contig of 2496 bp in length.

FEATURES  
source  
1..200000  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="4"  
1206..12031  
gap

gap /estimated\_length=unknown  
13179..24004  
/estimated\_length=unknown  
25126..35951  
/estimated\_length=unknown  
37176..48001  
/estimated\_length=unknown  
49548..60373  
/estimated\_length=unknown  
61457..72281  
/estimated\_length=unknown  
73416..84240  
/estimated\_length=unknown  
85494..96318  
/estimated\_length=unknown  
97556..108380  
/estimated\_length=unknown  
110361..121185  
/estimated\_length=unknown  
122779..133603  
/estimated\_length=unknown  
135589..146413  
/estimated\_length=unknown  
148102..158926  
/estimated\_length=unknown  
160583..171407  
/estimated\_length=unknown  
173904..184728  
/estimated\_length=unknown  
186680..197504  
/estimated\_length=unknown

ORIGIN

Query Match 1.7%; Score 52; DB 14; Length 200000;  
Best Local Similarity 100.0%; Pred. No. 5,7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCACTGCACCTCGAGCTGGGCAAGACAGCAAGACTCTGTCTC 3122  
Db 96734 CAAGATTGTGCACTGCACCTCGAGCTGGGCAAGACAGCAAGACTCTGTCTC 96683

RESULT 185  
CNS01DSQ  
LOCUS  
DEFINITION CNS01DSQ 200853 bp DNA linear PRI 26-SEP-2001  
Human chromosome 14 DNA sequence BAC R-747H7 of library RPCI-11  
from chromosome 14 of Homo sapiens (Human), complete sequence.

ACCESSION AL121839  
VERSION AL121839.3 GI:15796542  
KEYWORDS HTG.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.

REFERENCE 1 (bases 1 to 200853)  
AUTHORS Heilig,R., Petit,J.L., Vico,V., Dasilva,C., Robert,C., Wincker,P.,  
Brothier,P., Catolico,L., Barbe,V., Pelletier,E., Artiguenave,F.,  
Levy,M., Beckenbarg,R., Bruls,T., deBertoldi,V., Cruaud,C.,  
Geyrhofer,G., Saurin,W. and Weissenbach,J.  
TITLE Sequencing of the human chromosome 14  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 200853)  
AUTHORS Genoscope.  
TITLE Direct Submission  
JOURNAL Submitted (26-SEP-2001) Genoscope - Centre National de Sequencage :  
BP 191 91006 EVRY cedex - FRANCE (E-mail : seque@genoscope.cns.fr  
- Web : www.genoscope.cns.fr)  
COMMENT On Sep 27, 2001 this sequence version replaced gi:1955595.  
-----  
Center: Genoscope / Centre National de Sequencage  
Center code: GS  
Web site: http://www.genoscope.cns.fr/





complement (AL592078.10:44579..44621) ,  
complement (AL592078.10:43824..43914) ,  
complement (AL592078.10:42535..42682) ,  
complement (AL592078.10:42128..42273) ,  
complement (AL592078.10:41696..41850) ,  
complement (AL592078.10:39088..39236) )  
/gene="RP11-61114.3"  
/locus\_tag="RP11-61114.3-001"  
/standard\_name="OTTHUMP0000003575"  
/codon\_start=1  
/product="novel protein (DKZp434M202) "  
/protein\_id="CAH72850.1"  
/db\_xref="GI:55663063"  
/db\_xref="InterPro:IPR008191"  
/db\_xref="UniProt/TREMBL:Q5V219"  
/translation="MIQPRAPLVLEKSGSDESGFKTAIIQIAPAPVDLCETEKURA  
AFPAVLEKSGFLVLLRECFRDSWTALHSVGEAGLVTLSVTPKTFPMNAHVT  
BALHONMALFSTLAQAEQCYLEGSITMRCGTCLAEYHLDYGHANNRCVLDLRD  
TVAVMFIDFGQLATIPVOSLRSLSDSDFTTPTLOPFMLEKDIILSSVEYVHRLKKG  
KITGALNSAVTPASNTIAVNPRLPLGLCIQDA"  
/join(21773..22450,24648..24774,34530..34835,37203..37359,  
37824..37943,39551..44405)  
/gene="RP11-350G8.8"  
/locus\_tag="RP11-350G8.8-001"  
/join(21773..22450,24648..24774,34530..34835,37203..37359,  
37824..37943,39551..44405)  
/gene="RP11-350G8.8"  
/locus\_tag="RP11-350G8.8-001"  
/product="novel protein (LOC126669) "  
/note="match: ESTs: Em:AA256153.1 Em:AA331688.1  
Em:AA608795.1 Em:A1143934.1 Em:A1203653.1 Em:A128131.1  
Em:A1218256.1 Em:A1274859.1 Em:A1635642.1 Em:A1699294.1  
Em:A1797276.1 Em:A1824144.1 Em:BW014449.1 Em:BM991263.1  
Em:BX088810.1 Em:CA436491.1 Em:CA45011.1 Em:CB265391.1  
Em:CB270316.1 Em:CD300256.1 Em:CD359900.1 Em:CF552824.1  
Em:J139217.1  
match: cDNAs: Em:AK034662.1 Em:AK074067.1 Em:AK126006.1  
Em:AK131004.1"  
/join(21860..22450,24648..24774,34530..34835,37203..37359,  
37824..37943,39551..39737)  
/gene="RP11-350G8.8"  
/locus\_tag="RP11-350G8.8-001"  
/standard\_name="OTTHUMP00000034319"  
/note="match: proteins: Tr:QB8BD5 Tr:QB8E05"  
/codon\_start=1  
/product="novel protein (LOC126669) "  
/protein\_id="CAH72851.1"  
/db\_xref="GI:55663064"  
/db\_xref="GOA:Q5V218"  
/db\_xref="InterPro:IPR000980"  
/db\_xref="UniProt/TREMBL:Q5V218"  
/translation="NOMSEPTPQASACLCMASSLASSTAPTLIGRAGRPIMAAKPKF  
PLNLKTVERRAKPGQGGKLRKNSBEAGAGPKGRKNSAAELGSGAGVGPDSR  
LSRDSIQGLIOAAGGRKNSRATREBPATKSGSGSTYINRLIKVDYDKMKNR  
YSSSSSSSSSSSSASSPSSILPELDKIKIIOQETIILDDADPDYAKTKKQCD  
ARVBSNDGMEPYDAQOMITEIRRGSGDPLVKGDLSDCEBDDPKLSETLAKR  
BSKDLKGPOLYDTPYTPARAGPARAKRPPSRIPENDREAPAEAEQWEMKKE  
QIVRLSVQPEGAERSPREETVYRQHNRKSTQKLIKRLALSDHSEGEKVDGLPEK  
QWYHGALIRAEBSRLQPCKAGTIVRNSGNSRYSIALDSQGCWHIIVAQTKDN  
KTYLNGTSAVFISIPVYHYNSKLPFGAGAHMTLLVPHSKL"  
/join(37842..37943,53524..54114)  
/gene="RP11-350G8.8"  
/locus\_tag="RP11-350G8.8-002"  
/join(37842..37943,53524..54114)  
/gene="RP11-350G8.8"  
/locus\_tag="RP11-350G8.8-002"  
/product="novel protein (LOC126669) "  
/note="match: ESTs: Em:BM974630.1"  
44384..44389  
/gene="RP11-350G8.8"  
/locus\_tag="RP11-350G8.8-002"  
44405  
/gene="RP11-350G8.8"

gene  
/Locus\_tag="RP11-350G8.8-002"  
complement {join(54436..58752,69305..69398,73906..73975,  
75715..75761,87776..87917,88730..88896,89186..89367,  
93280..93403,94442..94690,118172..118693) )  
/gene="TL6R"  
/Locus\_tag="RP11-350G8.6-001"  
complement {join(54436..58752,69305..69398,73906..73975,  
75715..75761,87776..87917,88730..88896,89186..89367,  
93280..93403,94442..94690,118172..118693) )  
/gene="TL6R"  
/Locus\_tag="RP11-350G8.6-001"  
/product="interleukin 6 receptor"  
/note="match: ESTs: Em:AA035013.1 Em:AA193347.1  
Em:AA411559.1 Em:A1003332.1 Em:A134387.1 Em:A143504.1  
Em:A1971631.1 Em:AL120877.1 Em:AM177966.1 Em:AM178188.1  
Em:AM468643.1 Em:AM503691.1 Em:BM032730.1 Em:BM500539.1  
Em:BM575786.1 Em:BI018205.1 Em:BI760931.1 Em:BI860794.1  
Em:BM015654.1 Em:BM044442.1 Em:BM128112.1 Em:BM549585.1  
Em:BM911928.1 Em:BQ233533.1 Em:BQ335482.1 Em:BQ76559.1  
Em:BM997769.1 Em:BQ937176.1 Em:BU630287.1 Em:BU632645.1  
Em:BU848859.1 Em:BX113990.1 Em:BX509093.1 Em:BX646979.1  
Em:CA308138.1 Em:CA422880.1 Em:CA427041.1 Em:CB163591.1  
Em:CB962465.1 Em:CD516644.1 Em:CD609023.1 Em:CD609024.1  
Em:CD695637.1 Em:CF123018.1 Em:H59001.1  
match: cDNAs: Em:AF147881.1 Em:BC035327.1 Em:X12830.1  
Em:X58298.1"  
complement {54436)  
/gene="TL6R"  
/Locus\_tag="RP11-350G8.6-001"  
complement {54451..54456)  
/gene="TL6R"  
/Locus\_tag="RP11-350G8.6-001"  
complement {join(58506..58752,69305..69398,73906..73975,  
75715..75761,87776..87917,88730..88896,89186..89367,  
93280..93403,94442..94690,118172..118256) )  
/gene="TL6R"  
/Locus\_tag="RP11-350G8.6-001"  
/standard\_name="OTTHUMP00000034320"  
/note="match: proteins: Sw:P08887"  
/codon\_start=1  
/product="interleukin 6 receptor"  
/protein\_id="CAH72853.1"  
/db\_xref="GI:55663065"  
/db\_xref="InterPro:IPR002996"  
/db\_xref="InterPro:IPR003530"  
/db\_xref="InterPro:IPR003964"  
/db\_xref="InterPro:IPR007110"  
/db\_xref="InterPro:IPR008957"

Query Match 1.7% Score 52; DB 8; Length 201167;  
Best Local Similarity 10.0%; Pred. No. 5,76-16; indels 0;  
Matches 52; Conservative 0; Mismatches 0; Gaps 0;  
Qy 3071 CAAGATTGCGCACTGCATCCAGCTCGGGCAACAGCAAGACTGTGCTTC 3122  
122716 CAAGATTGCGCACTGCATCCAGCTCGGGCAACAGCAAGACTGTGCTTC 122767

RESULT 187  
AC090559 202539 bp DNA linear PRI 29-JUL-2002  
LOCUS Homo sapiens chromosome 11, clone RP11-750H9, complete sequence.  
DEFINITION AC090559  
ACCESSION AC090559.7 GI:22002211  
VERSION HTG.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens  
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 202539)  
AUTHORS Birren,B., Nusbaum,C. and Lander,B.  
TITLE Homo sapiens chromosome 11, clone RP11-750H9

JOURNAL  
REFERENCE  
AUTHORS  
Unpublished  
(bases 1 to 202539)  
Birren, B., Linton, L., Nusbaum, C., Lander, E., Allen, N., Anderson, S.,  
Barna, N., Bastien, V., Boguslavsky, L., Boukhgalter, B., Brown, A.,  
Cammarata, J., Campopiano, A., Choepel, Y., Colangelo, M., Collins, S.,  
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,  
Dodge, S., Fero, S., Ferreira, P., Fitzhugh, W., Gage, D., Galagan, J.,  
Gardyna, S., Glinde, S., Goyette, M., Graham, L., Grand-Pierre, N.,  
Hagos, B., Heatford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R.,  
Jones, C., Karatas, A., Labrocque, K., Lamazares, R., Landers, T.,  
Lehoczky, J., Levine, R., Liu, G., Maclean, C., Macdonald, P.,  
Marquis, N., Matthews, C., McCarthy, M., McSwan, P., McKernan, K.,  
McPheeters, R., Meldrim, J., Menus, L., Mihova, T., Mlenga, V.,  
Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C.H.,  
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R.,  
Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Roedel, M.,  
Roy, A., Santos, R., Schauer, S., Schuback, R., Seaman, S., Severy, P.,  
Souarez, C., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
Strauss, N., Subramanian, A., Talamas, J., Testfaye, S., Theodore, J.,  
Travers, M., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A.,  
Wilson, B., Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J.,  
Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
Direct Submission  
Submitted (03-MAR-2001) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 202539)  
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,  
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,  
Cammarata, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A.,  
Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,  
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,  
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A.,  
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,  
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,  
McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mlenga, V.,  
Murphy, T., Naylor, J., Nguyen, C., Nicoli, R., Norbu, C., Norman, C.H.,  
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,  
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,  
Testfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,  
Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
Direct Submission  
Submitted (26-JUN-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
4 (bases 1 to 202539)  
Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S.,  
Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhgalter, B.,  
Cammarata, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A.,  
Cooke, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S., Dodge, S.,  
Fero, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J.,  
Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B.,  
Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Karatas, A.,  
Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K.,  
Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,  
McCarthy, M., Meldrim, J., Menus, L., Mihova, T., Mlenga, V.,  
Murphy, T., Naylor, J., Nguyen, C., Nicoli, R., Norbu, C., Norman, C.H.,  
O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,  
Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P.,  
Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P.,  
Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J.,  
Testfaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,  
Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J.,  
Zembek, L., Zimmer, A. and Zody, M.

TITLE  
JOURNAL  
REFERENCE  
AUTHORS  
Direct Submission  
Submitted (29-JUL-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Jul 29, 2002 this sequence version replaced gi:21591871.  
All repeats were identified using RepeatMasker:  
Smit, A.F.A. & Green, P. (1996-1997)

FEATURES  
source  
http://ftp.genome.washington.edu/RW/RepeatMasker.html  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: MIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L11075  
Center clone name: 750\_H\_9  
-----  
Location/Qualifiers  
1..202539  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="11"  
/map="11"  
/clone="RP11-750H9"  
/clone\_1fb="RP11 Human Male BAC"  
996..1158  
/repeat\_family="CT-rich"  
2010..2190  
/rpt\_family="CCCCAG)n"  
/rpt\_family="ATUsq"  
3759..3937  
/rpt\_family="TTCA)n"  
3981..4135  
/rpt\_family="TTCA)n"  
complement(4149..4503)  
/rpt\_family="L1MA4A"  
4508..4785  
/rpt\_family="TTCA)n"  
4805..4903  
/rpt\_family="TTCA)n"  
complement(4911..4987)  
/rpt\_family="L2"  
5029..5059  
/rpt\_family="TTG)n"  
5521..5599  
/rpt\_family="TTCA)n"  
complement(5601..5959)  
/rpt\_family="L1PB3"  
5960..5987  
/rpt\_family="TA)n"  
complement(5988..6163)  
/rpt\_family="L1PB3"  
complement(6164..6470)  
/rpt\_family="ATUsq"  
complement(6473..6771)  
/rpt\_family="ATUsq"  
complement(6772..6832)  
/rpt\_family="L1PB3"  
complement(7731..7814)  
/rpt\_family="MIR"  
7841..8088  
/rpt\_family="GA-rich"  
complement(8333..8516)  
/rpt\_family="L1ME4A"  
8517..8827  
/rpt\_family="ATUsq"  
complement(8828..8939)  
/rpt\_family="L1ME4A"  
complement(8940..9144)  
/rpt\_family="L1MA3"  
complement(9145..9200)  
/rpt\_family="L1ME4A"  
complement(9194..9273)  
/rpt\_family="L1ME3A"  
9274..9580  
/rpt\_family="ATUsq"  
complement(9581..9641)  
/rpt\_family="L1ME3A"  
9642..9955  
/rpt\_family="ATUsq"

```

unsure      9915..9924
            /note="<30 qual SNGL region"
repeat_region complement(9956..10051)
            /rpt_family="LIME3A"
repeat_region complement(10052..10356)
            /rpt_family="AluSg"
repeat_region complement(10533..10835)
            /rpt_family="AluSx"
repeat_region complement(10878..10997)
            /rpt_family="Aluub"
repeat_region complement(10998..11149)
            /rpt_family="LIME3A"
repeat_region complement(11150..11184)
            /rpt_family="AluSx"
repeat_region 11185..11208
            /rpt_family="TTTTA)n"
repeat_region complement(11209..11495)
            /rpt_family="AluSx"
repeat_region complement(11500..11622)
            /rpt_family="FLAM_C"
repeat_region 11667..11805
            /rpt_family="Aluub"
repeat_region 11806..11830
            /rpt_family="Aluub"
repeat_region 11831..11976
            /rpt_family="TTGG)n"
repeat_region 12021..12326
            /rpt_family="AluSg"
repeat_region complement(12952..13398)
            /rpt_family="L2"
repeat_region complement(13452..13611)
            /rpt_family="MIR"
repeat_region complement(13636..13765)
            /rpt_family="MIR"
repeat_region 14462..14661
            /rpt_family="MIR"
repeat_region 14692..14870
            /rpt_family="MIR"
repeat_region complement(15253..15394)
            /rpt_family="MIR"
repeat_region complement(15871..15976)

Query Match      1.7%: Score 52; DB 8; Length 202539;
Best Local Similarity 100.0%; Pred. No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2889 GAGGCGAGTGTGATCCTGAGGCGAGGCTTGAGACGAGCCTGCGCAACAT 2940
        |||||
Db      63712 GAGGCGAGTGTGATCCTGAGGCGAGGCTTGAGACGAGCCTGCGCAACAT 63763

RESULT 188
AC151042      202971 bp      DNA      linear      HTG 24-NOV-2004
LOCUS      Callithrix jacchus clone CH259-63D17, WORKING DRAFT SEQUENCE, 15
DEFINITION      ordered pieces.
AC151042      GI:56044586
VERSION      HTG; HTGS_PHASE2; HTGS_DRAFT.
KEYWORDS      Callithrix jacchus (white-tufted-ear marmoset)
SOURCE      Callithrix jacchus
ORGANISM      Buthryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Platyrrhini;
            Callitrichidae; Callithrix.
            1 (bases 1 to 202971)
REFERENCE      Antonellis,A., Ayale,K., Benjamin,B., Blakesley,R.W., Boakye,A.,
            Bouffard,G.G., Brinkley,C., Brooks,S., Chu,G., Coleman,H.,
            Engle,J., Guan,X., Gupta,J., Haghighi,P., Han,J., Hansen,N.,
            Ho,S.-L., Hu,P., Hurle,B., Idol,J.R., Jones,C., Kwong,P., Latic,P.,
            Larson,S., Lee-Lin,S.-Q., Legaaspi,R., Madden,M., Maduro,Q.L.,
            Maduro,V.B., Margulies,E.H., Mastello,C., Maskeri,B., McDowell,J.,
            Mullikin,J.C., Oestreichner,J.S., Park,M., Portnoy,M.E., Prasad,A.,
            Puri,O., Reddik-Dugue,N., Rosas,B., Schandler,K., Schueler,M.G.,

```

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

Shah,K., Sison,C., Stantirpop,S., Stephen,B., Thomas,J.W.,  
 Thomas,P.J., Telpouri,V., Vogt,J.L., Wehenby,K.D., Young,A. and  
 Green,B.D.  
 NISC Comparative Sequencing Initiative  
 Unpublished  
 2 (bases 1 to 202971)  
 Direct Submission  
 Green,B.D.  
 Submitted (19-AUG-2004) NIH Intramural Sequencing Center, 8717  
 Grovemont Circle, Gaithersburg, MD 20877, USA  
 3 (bases 1 to 202971)  
 Direct Submission  
 Green,B.D.  
 Submitted (24-NOV-2004) NIH Intramural Sequencing Center, 5625  
 Fishers Lane, Rockville, MD 20852, USA  
 On Nov 24, 2004 this sequence version replaced gi:51372024.  
 ----- Genome Center  
 Center: NIH Intramural Sequencing Center  
 Center code: NISC  
 Web site: http://www.nisc.nih.gov  
 Contact: nisc\_zoo@ngri.nih.gov  
 ----- Project Information  
 Center project name: hja  
 Center clone name: 063D17

The sequence data in this record represents an 'enhanced'  
 version of a Phase 2 submission. Specifically, the indicated  
 order and orientation of each sequence contig has been  
 established using one or more of the following: read-pair  
 data from individual subclones, overlaps with neighboring  
 clones, alignment with available reference sequence (e.g.,  
 human), and/or confirmation by PCR testing. In addition,  
 the sequence assembly is based on at least 8X average  
 coverage in Q20 bases and has been revised to rule out  
 gross misassemblies, the low-quality ends of sequence  
 contigs have been trimmed away, and each base is associated  
 with a Phrap-derived quality score.

#### ----- Summary Statistics

Sequencing vector: plasmid; n/a; 100% of reads  
 Chemistry: Dye-terminator Big Dye; 100% of reads  
 Assembly program: Phrap; version 0.990319  
 Consensus quality: 199626 bases at least Q40  
 Consensus quality: 200659 bases at least Q30  
 Consensus quality: 201280 bases at least Q20  
 Insert size: 22800; agarose-fp  
 Insert size: 201571; sum-of-contigs  
 Quality coverage: 7.60x in Q20 bases; agarose-fp  
 Quality coverage: 8.60x in Q20 bases; sum-of-contigs

\* NOTE: This is a 'working draft' sequence. It currently  
 \* consists of 15 contigs. Gaps between the contigs  
 \* are represented as runs of N. The order of the pieces  
 \* is believed to be correct as given, however the sizes  
 \* of the gaps between them are based on estimates that have  
 \* been provided by the submitter.  
 \* This sequence will be replaced  
 \* by the finished sequence as soon as it is available and  
 \* the accession number will be preserved.  
 \* 1 19090: contig of 19090 bp in length  
 \* 19091 19190: gap of unknown length  
 \* 19191 32965: contig of 13775 bp in length  
 \* 32966 33065: gap of unknown length  
 \* 33066 37634: contig of 4569 bp in length  
 \* 37635 37734: gap of unknown length  
 \* 37735 39153: contig of 1419 bp in length  
 \* 39154 39253: gap of unknown length  
 \* 39254 46148: contig of 6895 bp in length  
 \* 46149 46248: gap of unknown length  
 \* 46249 75331: contig of 29083 bp in length  
 \* 75332 75431: gap of unknown length  
 \* 75432 79016: contig of 3585 bp in length  
 \* 79017 79116: gap of unknown length  
 \* 79117 82235: contig of 3119 bp in length

```
*      82236      82335: gap of unknown length
*      82336      87728: contig of 5393 bp in length
*      87728      87828: gap of unknown length
*      87828      100513: contig of 12685 bp in length
*      100513      100614: gap of unknown length
*      100614      123670: contig of 23057 bp in length
*      123670      123771: gap of unknown length
*      123771      137796: contig of 14026 bp in length
*      137796      137896: gap of unknown length
*      137896      161136: contig of 23240 bp in length
*      161136      161236: gap of unknown length
*      161236      176590: contig of 15354 bp in length
*      176590      176691: gap of unknown length
*      176691      202971: contig of 26281 bp in length.
FEATURES
source      Location/Qualifiers
            1..202971
            /organism="Callithrix jacchus"
            /mol_type="genomic DNA"
            /db_xref="taxon:9483"
            /clone="CH259-63D17"
            /clone_1lb="CH259"
            /note="BAC resource: http://bacpac.chori.org/"
misc_feature 1..39045
            /note="clone overlaps with GenBank Accession Number
            AC151015 clone CH259-164B19 (center project name h1z)"
misc_feature 1..19090
            /note="assembly_fragment
            clone_end:17
            vector_side:left"
gap          19091..19190
            /estimated_length=unknown
            /estimated_length=unknown
            /estimated_length=unknown
            /note="assembly_fragment"
gap          32966..33065
            /estimated_length=unknown
            /note="assembly_fragment"
misc_feature 33066..33734
            /note="assembly_fragment"
            /estimated_length=unknown
            /estimated_length=unknown
            /estimated_length=unknown
            /note="assembly_fragment"
misc_feature 37735..39153
            /note="assembly_fragment"
            /estimated_length=unknown
            /estimated_length=unknown
            /estimated_length=unknown
            /note="assembly_fragment"
gap          39154..39253
            /estimated_length=unknown
            /estimated_length=unknown
            /estimated_length=unknown
            /note="assembly_fragment"
misc_feature 46149..46248
            /estimated_length=unknown
            /estimated_length=unknown
            /estimated_length=unknown
            /note="assembly_fragment"
gap          46249..75331
            /estimated_length=unknown
            /estimated_length=unknown
            /estimated_length=unknown
            /note="assembly_fragment"
gap          75332..75431
            /estimated_length=unknown
            /estimated_length=unknown
            /estimated_length=unknown
            /note="assembly_fragment"
misc_feature 75432..79016
            /note="assembly_fragment"
            /estimated_length=unknown
            /estimated_length=unknown
            /estimated_length=unknown
            /note="assembly_fragment"
gap          79017..79116
            /estimated_length=unknown
            /estimated_length=unknown
            /estimated_length=unknown
            /note="assembly_fragment"
misc_feature 79117..82235
            /note="assembly_fragment"
            /estimated_length=unknown
            /estimated_length=unknown
            /estimated_length=unknown
            /note="assembly_fragment"
gap          82236..82335
            /estimated_length=unknown
            /estimated_length=unknown
            /estimated_length=unknown
            /note="assembly_fragment"
misc_feature 82336..87728
            /note="assembly_fragment"
            /estimated_length=unknown
            /estimated_length=unknown
            /estimated_length=unknown
            /note="assembly_fragment"
gap          87729..87828
            /estimated_length=unknown
            /estimated_length=unknown
            /estimated_length=unknown
            /note="assembly_fragment"
misc_feature 87829..100513
            /note="assembly_fragment"
            /estimated_length=unknown
            /estimated_length=unknown
            /estimated_length=unknown
            /note="assembly_fragment"
misc_feature 98925..202971
            /note="clone overlaps with GenBank Accession Number
            AC150824 clone CH259-392G13 (center project name h1b)"
gap          100514..100613
            /estimated_length=unknown
            /estimated_length=unknown
            /estimated_length=unknown
            /note="assembly_fragment"
misc_feature 100614..123670
            /estimated_length=unknown
            /estimated_length=unknown
            /estimated_length=unknown
            /note="assembly_fragment"
gap          123671..123770
```

```
misc_feature /estimated_length=unknown
            123771..137796
            /note="assembly_fragment"
gap          137797..137896
            /estimated_length=unknown
            /estimated_length=unknown
            /estimated_length=unknown
            /note="assembly_fragment"
misc_feature 137897..161136
            /note="assembly_fragment"
            161137..161236
            /estimated_length=unknown
            /estimated_length=unknown
            /estimated_length=unknown
            /note="assembly_fragment"
misc_feature 161237..176590
            /note="assembly_fragment"
            176591..176690
            /estimated_length=unknown
            /estimated_length=unknown
            /estimated_length=unknown
            /note="assembly_fragment"
misc_feature 176691..202971
            /note="assembly_fragment
            clone_end:SP6
            vector_side:right"
ORIGIN
Query Match      1.7%; Score 52; DB 14; Length 202971;
Best Local Similarity 100.0%; Pred. No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      3071 CAAGATTGTGCCACTGCACTGCCAGCTGGCAACAGAGCAAGACTGTCTC 3122
Db      103406 CAAGATTGTGCCACTGCACTGCCAGCTGGCAACAGAGCAAGACTGTCTC 103457
RESULT 189
AC008040/c      204917 bp      DNA      linear      PRI 28-SEP-2002
LOCUS      Homo sapiens 3 BAC RP11-379K17 (Roswell Park Cancer Institute Human
DEFINITION      BAC library) complete sequence.
ACCESSION      AC008040
VERSION      AC008040.7 GI:5922025
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 204917)
Muzny,D.M., Adams,C., Bailey,M., Barbrie,J., Blankenburg,K.,
Bodota,B., Bouck,J., Bowie,S., Brooks,A., Buhay,C., Bunac,C.,
Burkelt,C., Burrows,J., Carter,M., Chacko,J., Chen,Z., Cox,C.,
David,R., Delgado,O., Deshazo,D., Ding,Y., Doman-Rashtid,N.,
Dugan-Rocha,S., Durbin,K.J., Fernandez,C., Ferraguto,D.,
Forcum-Tansey,J., Frantz,P., Ganesh,R., Garcia,D.K., Gorrell,J.H.,
Gorrell,L.L., Guevara,W., Harris,K., He,X., Hernandez,J.,
Hodgson,A., Hognes,M., Holloway,C., Hosak,H., Jackson,L.B.,
Jackson,L., Jia,Y., Jones,M., Kelly,S., Kondajewski,N., Kong,Y.,
Kovar,C., Leal,B., Li,Z., Lichareye,O., Liu,J., Liu,W., Logan,O.,
Lu,J., Lucier,R., Martin,R., Martinez,C., McLeod,M.P., Mei,G.,
Moore,S., Moorish,T., Morgan,M., Morris,S., Nash,S., Nelson,A.,
Nguyen,R., Nguyen,N., Nguyen,S., Oswal,G., Parish,B., Paxton,S.,
Payton,B., Perez,L., Pu,L.L., Qulles,M., Reiter,D., Rives,M.,
Samuel,S., Say,J., Scherer,S., Shah,E., Shen,H., Simon,M.,
Sparks,A., Stamps,A., Sugang,R., Tabors,P., Taylor,T., Vaequez,L.,
Vinson,R., Vo,Q., Wabhan,M., Watlington,S., Weinstein,G.,
Weinstock,I.R., Williams,A., Worley,K., Wren,J., Wrenford,G.,
Yu,W., Zhou,X., Naylor,S.L., Nelson,D. and Gibbs,R.
Direct Submission
Unpublished
2 (bases 1 to 204917)
Worley,K.C.
Direct Submission
Submitted (14-JUL-1999) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 204917)
Worley,K.C.
Direct Submission
Submitted (24-SEP-1999) Human Genome Sequencing Center, Department
```

of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
4 (baaes 1 to 204917)  
Worley, K.C.  
Direct Submision  
Submitted (09-OCT-1999) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
5 (baaes 1 to 204917)  
Worley, K.C.  
Direct Submision  
Submitted (16-MAY-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
6 (baaes 1 to 204917)  
Worley, K.C.  
Direct Submision  
Submitted (28-JUN-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
7 (baaes 1 to 204917)  
Worley, K.C.  
Direct Submision  
Submitted (28-SEP-2002) Human Genome Sequencing Center, Department of Molecular and Human Genetics, Baylor College of Medicine, One Baylor Plaza, Houston, TX 77030, USA

On Sep 24, 1999 this sequence version replaced gi:573555.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email [gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

COMMENT

CLONE LENGTH: This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features listing.

#### ANNOTATION OF FEATURES:

STS are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green, unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 2 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as low coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality standards - estimated error rate less than 1 per 10,000 bases. Reports of lowest quality individual bases and measures of base quality are listed below. Description of the metrics can be found at URL:

<http://gc.bcm.tmc.edu:8088/quality.info/genbank.annotation.html>.

#### QUALSTAT-REPORT.

FEATURES

source

Location/Qualifiers

1..204917

/organism="Homo sapiens"

/mol\_type="genomic DNA"

/db\_xref="taxon:9606"

/chromosome="3"

/clone="RP11-379K17"

misc\_feature

complement(1..1995)

/note="Overlaps bases 1..1995 of clone AC023891"  
/function="clone overlap"  
766..1132  
/note="Region: od1d01.st Homo sapiens cDNA, AA810014"  
misc\_feature  
2456..2279  
/note="Region: z855f12.st Homo sapiens cDNA, AA286777"  
STS  
2554..2768  
/standard\_name="G37704"  
/db\_xref="dbSTS:55264"  
repeat\_region  
6897..7126  
/rpt\_family="MER20"  
misc\_feature  
7488..8243  
/note="Region: Unigene cluster containing AA400519 and AA165136"  
repeat\_region  
11171..11471  
/rpt\_family="AluX"  
complement(13526..13835)  
repeat\_region  
/rpt\_family="AluSg"  
misc\_feature  
13556..15048  
/note="Region: Unigene cluster containing A1683116, A1048825, AA278956, and M38986"  
repeat\_region  
15264..15890  
/rpt\_family="AluX"  
complement(14286..14586)  
misc\_feature  
/note="Region: Unigene cluster containing A1042019 and A1890489"  
repeat\_region  
15800..15834  
/rpt\_family="AT-rich"  
repeat\_region  
16816..16837  
/rpt\_family="AT-rich"  
repeat\_region  
16911..17154  
/rpt\_family="MIR"  
complement(17196..17318)  
repeat\_region  
/rpt\_family="MER112"  
17598..17782  
/rpt\_family="MER20"  
misc\_feature  
18072..19357  
/note="Region: Unigene cluster containing A1767291, A1522162, M31729 and AA554506"  
STS  
18120..18226  
/standard\_name="G24029"  
/db\_xref="dbSTS:37867"  
20464..20496  
/rpt\_family="(CAAAA)n"  
repeat\_region  
20576..20882  
/rpt\_family="AluG"  
complement(20890..21061)  
repeat\_region  
/rpt\_family="MER46A"  
21460..21885  
/rpt\_family="L1PA13"  
repeat\_region  
22470..22760  
/rpt\_family="AluB"  
complement(24017..24179)  
repeat\_region  
/rpt\_family="Charlieb"  
complement(24180..24635)  
repeat\_region  
/rpt\_family="L1R22A"  
complement(24639..24744)  
repeat\_region  
/rpt\_family="HERVK22"  
complement(24750..25603)  
repeat\_region  
/rpt\_family="HERVK22"  
complement(25606..26058)  
repeat\_region  
/rpt\_family="L1R22A"  
complement(26061..26174)  
repeat\_region  
/rpt\_family="Charlieb"  
complement(26391..26559)  
repeat\_region  
/rpt\_family="MIR"  
26635..26943  
/rpt\_family="AluX"  
repeat\_region  
27836..28013  
/rpt\_family="MIR11"  
complement(28215..28526)  
repeat\_region  
/rpt\_family="AluSg"  
28548..28687  
repeat\_region

```
repeat_region /rpt_family="LM4"
28688 .28981
/rpt_family="AluB"
repeat_region 28982 .29281
/rpt_family="LM4"
repeat_region 29282 .29588
/rpt_family="AluSx"
repeat_region 29589 .30042
/rpt_family="LM4"
repeat_region 30055 .30213
/rpt_family="LM4"
```

Query Match	1.7%;	Score 52;	DB 8;	Length 204917;
Best Local Similarity	100.0%;	Pred. No. 5.7e-16;		
Matches 52;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY	3071	CAAGATTGTGCCACTGCACTCAGCCTGGGGCAACAGAGCAAGACTCTGTCTC	3122
Db	160446	CAAGATTGTGCCACTGCACTCAGCCTGGGGCAACAGAGCAAGACTCTGTCTC	160395

RESULT	190				
	AC073916/c				
LOCUS	AC073916	205283 bp	DNA	linear	FRI 27-MAR-2007
DEFINITION	Homo sapiens 12 BAC RP11-408I18 (Roswell Park Cancer Institute Human BAC Library) complete sequence.				

ACCESSION	AC073916	GI:29293998
VERSION	AC073916.41	
KEYWORDS	HTG.	
SOURCE	Homo sapiens	(human)
ORGANISM	Homo sapiens	

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.

## REFERENCE AUTHORS

Muzny, D.M., Adams, C., Abo-Okola, B., Al-Osman, F.R., Allen, C., Allbrooks, S.L., Amarantunga, H.C., Are, J.R., Ayale, M., Banks, T., Barbarella, J., Benton, J., Blinage, K., Blankenburg, K., Bonini, D., Bouck, U., Bowle, S., Brileva, M., Brown, E., Brown, M., Bryant, N.P., Buhay, C., Burch, P., Burrett, C., Butwell, K.L., Byrd, N.C., Cartron, T.F., Carter, M., Cavazos, S.R., Chacko, J., Chavez, D., Chen, G., Chen, R., Chen, Z., Chu, D., Chowdhry, I., Christopoulos, C., Cleveland, C.D., Cox, C., Coyle, M.D., Datnome, S.R., David, R., Davila, M.L., Davis, C., Davy-Carroll, L., Dederich, D.A., Delaney, K.R., Delgado, C., Dem, A.L., Ding, Y., Dinh, H.H., Douthwaite, K.J., Draper, H., Dugan-Rocha, S., Durbin, K.J., Earnhart, C., Edgar, D., Edwards, C.C., Elhaj, C., Emmerling, S., Escotto, M., Falla, T., Fargueto, D., Flagg, N., Ford, J., Foster, P., Frantz, P., Gafisi, A., Gao, J., Garcia, A., Garner, T., Garra, N., Gill, R., Gorrell, J.H., Guetara, W., Gunaratne, P., Hale, S., Hamilton, K., Han, J., Harris, C., Harris, K., Hart, M., Havlik, P., Hawes, A., Hernandez, J., Hernandez, O., Hodgson, A., Hogue, W., Holloway, C., Hollins, B., Homsi, F., Howard, S., Huber, J., Huiyk, S., Hume, J., Ioshikhes, I., Jackson, L.E., Jacobson, B., Jia, Y., Johnson, R., Jolyvet, S., Joudah, S., Karlsson, E., Kelly, S., Khan, U., King, L., Korvah, J., Kovar, C., Kravtovic, J., Kureshi, A., Landry, N., Leal, B., Lee, E., Lewis, L.C., Lewis, L., Li, Z., Li, Z., Licharge, O., Lien, C., Liu, Y., Liu, W., Loubegeed, H., Lozdo, R.J., Lu, X., Lueker, A., Lucier, R., Luna, R., Ma, J., Maheshwari, M., Mapua, P., Marandei, I., Martin, R., Matindale, A., Martinez, E., Massey, E., Maxwell, E., McLeod, M.P., Meador, M., Mei, G., Merscher, S., Metker, M., Miller, A., Miner, G., Miner, Z., Mitchell, S., Mohabbat, K., Montgomery, K.T., Morgan, M., Morris, S., Moser, M., Mohabbat, K., Nelson, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N., Nguyen, N., Nickerson, E., Nookemko, S., Oguni, M., Okumura, G., Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L., Peters, L., Pickens, R., Pirbams, E., Pu, L.L., Quiles, M., Ren, Y., Rivers, M., Rojka, A., Rojchubank, I., Rolfe, M., Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shim, C., Shooshbari, N., Sisson, I., Sodergren, E., Sonalle, T., Sparks, A., Stanley, H., Stone, H., Sutton, A., Svatek, A., Taboc, P., Tamezias, A., Tamezias, K., Tang, H., Tamez, J., Taylor, C., Taylor, T., Telitold, B., Thomas, R., Thomas, S., Tsemali, K., Tvaquer, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,

TITLE	Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watling, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K., Wu, C., Wu, Y., Wu, Y., Zhou, J., Zorrilla, S., Kuchelapatti, R., Weinstock, G. and Gibbs, R.
JOURNAL	Direct Submission
REFERENCE	Unpublished
AUTHORS	2 (bases 1 to 205283)
TITLE	Worley, K. C.
JOURNAL	Direct Submission
	Submitted (05-JUL-2000) Human Genome Sequencing Center, Department

REFERENCE 3 (bases 1 to 205283)  
AUTHORS Worley, K.C.  
TITLE Direct Submision  
JOURNAL Submitted (31-JUL-2002) Human Genome Sequencing Center, Department

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
4 (bases 1 to 205283)  
Worley, K. C.  
Direct Submission  
Submitted (01-AUG-2002) Human Genome Sequencing Center, Department

REFERENCE 5 (bases 1 to 205283)  
AUTHORS Worley, K.C.  
TITLE Direct Submission  
JOURNAL Submitted (06-MAR-2003) Human Genome Sequencing Center, Department

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
Submitted (15-MAR-2003)

6 (bases 1 to 205283)  
Worley, K. C.  
Direct Submission  
Human Genome Sequencing Center, Department

REFERENCE  
AUTHORS  
TITLE  
JOURNAL

7 (bases 1 to 205283)  
Worley, K. C.  
Direct Submission  
Submitted (27-MAR-2003) Human Genome Sequencing Center, Department  
of Molecular and Cellular Biology, University of Texas at Austin, Austin, TX 78712-1083, USA  
E-mail: worley@mcdb.utexas.edu

COMMENT On Mar 27, 2003 this sequence version replaced gi:22038291.  
INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email  
[gc-help@bcm.tmc.edu](mailto:gc-help@bcm.tmc.edu)

**CLONE LENGTH:** This sequence does not necessarily represent the entire insert of this clone. Overlapping regions of clones are only sequenced and submitted once, so the sequence for the remainder of the insert may be found in the record for the adjacent clones. Overlapping clones are noted at the beginning and end of the Features Listing.

**ANNOTATION *F* FEATURES:** STRs are identified using ePCR (Genome Res. 7:541-550) searches of a local database that includes entries from dbSTS, GDB, and local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green unpublished.) for Human and Mouse sequences.

Genes and Region of sequence similarity are identified by BLAST (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-36) to the EST and cDNA sequences. Genes demonstrate at least two exons flanked by consensus splice sites that maintained sequence continuity across the splice junctions. Sequences that are not identical matches are annotated as similar.

SEQUENCING LEAD COVERAGE: sequencing is completed to a minimum standard of double strand coverage with a minimum of 2 clones and 3 reads with no ambiguities or 2 chemistries with a minimum of 2 clones and 3 reads with no ambiguities. If the sequence quality for a region does not meet this standard, it will be indicated in the annotation as Low Coverage.



FEATURES	Location/Qualifiers
source	1. .205283

```

482_856
/strand_name="D125820"
/rpc_family="AT_rich"
710_845
/rpc_family="(TCCA)n"
complement(846_980)
/rpc_family="L2"
/rpc_family="L2"
1320_1681
/rpc_family="L2"
2124_2158
/rpc_family="AT_rich"
2327_2353
/rpc_family="(T)n"
complement(2835_2214)
/rpc_family="MERS8B"
3015_3377
/rpc_family="L1MC4"
complement(3384_3682)
/rpc_family="LTR40a"
3908_3944
/rpc_family="U2"
3983_4182
/rpc_family="L1MC4"
4217_4256
/rpc_family="(GA)n"
5472_5647
/rpc_family="(TG)n"
5698_5870
/rpc_family="(TG)n"
complement(6009_6188)
/rpc_family="MIR"
6410_6484
/rpc_family="MIRb"
complement(6501_6615)
/rpc_family="AlusX"
7401_7603
/rpc_family="L2"
8096_8186
/rpc_family="MERS96B"
complement(8187_8485)
/rpc_family="AlusX"
8486_8515
/rpc_family="MERS96B"
9052_9174
/rpc_family="L2"
10287_10476
/rpc_family="(CGGG)n"
10415_10551
/rpc_family="G-rich"
10603_10631
/rpc_family="GC_rich"
10658_10684
/rpc_family="GC_rich"
10781_10801

```

```

/rpt_family="(CCCCAG)n"
10835. .10859
/rpt_family="GC_rich"
11332. .11465

```

```
1.74; Score 52; DB 8; length 205283;
100.0%; Pred. No. 5.7e-16;
div 0; Mismatches 0; Indels 0; Gaps 0;
```

3071 CAAGTTGTGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
|||||  
82261 CAAGATTGTGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 82910

## RESULT 191

LOCUS	AC023560	205736 bp	DNA	linear	HTG 12-MAR-2000
DEFINITION	Homo sapiens chromosome 17 clone RP11-85120 map 17, WORKING DRAFT				
SEQUENCE	SEQUENCE, 34 unordered pieces.				

VERSION AC023560.2 GI:7229913  
KEYWORDS HTG; HTGS\_PHASE1; HTGS\_DRAFT.

## ORGANISM

REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE
1	Bairren, B., Linton, L., Nusbaum, C. and Lander, E.	Homo sapiens chromosome 17, clone RP11-65120	Unpublished	2
2	Bairren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.			

TITLE  
 JOURNAL  
 COMMENT

Anderson, S., Baldwin, J., Barna, N., Beda, F., Boguslavsky, L.,  
 Boukhalter, B., Brown, A., Burkett, G., Campobello, A., Castle, A.,  
 Choemel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,  
 Dearlano, K., Dewar, K., Dodge, S., Domingo, M., Doyle, M.,  
 Feneator, J., Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D.,  
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
 Grand-Pierre, N., Grant, G., Hages, B., Heaford, A., Horton, L.,  
 Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kam, L., Karakas, A.,  
 Klein, J., Landers, T., Laroque, K., Lehoczy, T., Levine, R.,  
 Lieu, C., Liu, G., Locke, K., MacDonald, P., Marguis, N., McCarthy, M.,  
 McEwen, P., McGuirk, A., McKernan, K., McPheters, R., Meldrum, J.,  
 Menous, L., Mihova, T., Miranda, C., Menga, V., Morrow, T., Naylor, J.,  
 Norman, C.H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, T.M.,  
 Peterson, K., Pierre, N., Plesni, C., Pollara, V., Raymond, C.,  
 Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S.,  
 Seavey, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,  
 Sudrmanhan, A., Talamas, J., Tefsey, S., Theodore, J., Tirrell, A.,  
 Travers, M., Triggillo, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B.,  
 Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and  
 Zody, M.

Direct Submission  
 Submitted (15-FEB-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Mar 12, 2000 this sequence version replaced 616978256.

p://ftp.genome.washington.edu/RM/RepeatMasker.html  
 t, A.E.A. & Green, P. (1996-1997)  
 repeats were identified using RepeatMasker:  
 Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: WIRB  
 Web site: <http://www-seq.wi.mit.edu>  
 Contact: [sequence\\_submissions@genome.wi.mit.edu](mailto:sequence_submissions@genome.wi.mit.edu)  
 Project Information  
 Center project name: L6741  
 Center clone name: 85.I.20  
 Summary Statistics  
 Sequencing vector: M13, M77815, 100% of reads  
 Chemistry: Dye-terminator Big Dye, 100% of reads  
 Assembly program: Phrap, Version 0.960731  
 Consensus quality: 17866 bases at least Q40

Consensus quality: 190940 bases at least Q30  
Consensus quality: 197088 bases at least Q20  
Insert size: 194000; agarose-fp  
Insert size: 202436; sum-of-contigs  
Quality coverage: 4.2 in Q20 bases; agarose-fp  
Quality coverage: 4.0 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 34 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 1267: contig of 1267 bp in length  
1268 1367: gap of 100 bp  
1368 2402: contig of 1035 bp in length  
2403 2502: gap of 100 bp  
2503 3524: contig of 1022 bp in length  
3525 3624: gap of 100 bp  
3625 4913: contig of 1289 bp in length  
4914 5013: gap of 100 bp  
5014 6289: contig of 1276 bp in length  
6290 6390: gap of 100 bp  
6390 7655: contig of 1266 bp in length  
7656 7755: gap of 100 bp  
7756 9267: contig of 1512 bp in length  
9268 9367: gap of 100 bp  
9368 11138: contig of 1771 bp in length  
11139 13822: contig of 2584 bp in length  
13823 13922: gap of 100 bp  
13923 16968: contig of 3046 bp in length  
16969 17068: gap of 100 bp  
17069 20185: contig of 3117 bp in length  
20186 20285: gap of 100 bp  
20286 22634: contig of 2349 bp in length  
22635 22734: gap of 100 bp  
22735 27012: contig of 4278 bp in length  
27013 27112: gap of 100 bp  
27113 29342: contig of 2230 bp in length  
29343 29442: gap of 100 bp  
29443 33381: contig of 3939 bp in length  
33382 33481: gap of 100 bp  
33482 36662: contig of 3081 bp in length  
36663 36662: gap of 100 bp  
36663 40982: contig of 4320 bp in length  
40983 41082: gap of 100 bp  
41083 45520: contig of 4438 bp in length  
45521 45620: gap of 100 bp  
45621 50623: contig of 5003 bp in length  
50624 50723: gap of 100 bp  
50724 55610: contig of 4881 bp in length  
55611 55710: gap of 100 bp  
55711 61038: contig of 5338 bp in length  
61039 61138: gap of 100 bp  
61139 66159: contig of 5021 bp in length  
66160 66259: gap of 100 bp  
66260 73479: contig of 7220 bp in length  
73480 73579: gap of 100 bp  
73580 80554: contig of 6975 bp in length  
80555 80654: gap of 100 bp  
80655 89313: contig of 8659 bp in length  
89314 89413: gap of 100 bp  
89414 98132: contig of 8719 bp in length  
98133 98232: gap of 100 bp  
98233 107469: contig of 9237 bp in length  
107470 107569: gap of 100 bp  
107570 117873: contig of 10304 bp in length  
117874 117973: gap of 100 bp  
117974 130973: contig of 13000 bp in length  
130974 131073: gap of 100 bp

\* 131074 143236: contig of 12163 bp in length  
\* 143237 143336: gap of 100 bp  
\* 143337 154478: contig of 11142 bp in length  
\* 154479 154578: gap of 100 bp  
\* 154579 169967: contig of 13389 bp in length  
\* 169968 170067: gap of 100 bp  
\* 170068 187347: contig of 17280 bp in length  
\* 187348 187447: gap of 100 bp  
\* 187448 205736: contig of 18289 bp in length.  
FEATURES  
source  
1. .205736  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="17"  
/map="17"  
/clone="RP11-85120"  
/clone\_11b="RP11 Human Male BAC"  
1. .1267  
/note="assembly\_fragment"  
1268. .1367  
/estimated\_length=100  
1368. .2402  
/note="assembly\_fragment"  
2403. .2502  
/estimated\_length=100  
2503. .3524  
/note="assembly\_fragment"  
3525. .3624  
/estimated\_length=100  
3625. .4913  
/note="assembly\_fragment"  
4914. .5013  
/estimated\_length=100  
5014. .6289  
/note="assembly\_fragment"  
6290. .6389  
/estimated\_length=100  
6390. .7655  
/note="assembly\_fragment"  
7656. .7755  
/estimated\_length=100  
7756. .9267  
/note="assembly\_fragment"  
9268. .9367  
/estimated\_length=100  
9368. .11138  
/note="assembly\_fragment"  
11139. .11238  
/estimated\_length=100  
11239. .13822  
/note="assembly\_fragment"  
13823. .13922  
/estimated\_length=100  
13923. .16968  
/note="assembly\_fragment"  
16969. .17068  
/estimated\_length=100  
17069. .20185  
/note="assembly\_fragment"  
20186. .20285  
/estimated\_length=100  
20286. .22634  
/note="assembly\_fragment"  
22635. .22734  
/estimated\_length=100  
22735. .27012  
/note="assembly\_fragment"  
27013. .27112  
/estimated\_length=100

Query Match 1.7%; Score 52; DB 14; Length 205736;  
Best Local Similarity 100.0%; Pred. No. 5,7e-16;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTCACCTGCACTCCAGCTTGCGCAACAGACAGACTCTGTCTC 3122  
 115864 CAAGATTGTCACCTGCACTCCAGCTTGCGCAACAGACAGACTCTGTCTC 115915

RESULT 192  
 AC015884/c 207661 bp DNA linear PRI 27-OCT-2002  
 LOCUS AC015884 Homo sapiens chromosome 17, clone RP11-216P6, complete sequence.  
 DEFINITION AC015884  
 AC015884.15 GI:24415419  
 VERSION HTG.  
 KEYWORDS Homo sapiens (human)  
 SOURCE Homo sapiens  
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 207661)  
 Birren, B., Nussbaum, C., and Lander, E.  
 Homo sapiens chromosome 17, clone RP11-216P6  
 Unpublished  
 2 (bases 1 to 207661)  
 Birren, B., Linton, L., Nussbaum, C., Lander, E., Allen, N., Anderson, M., Baldwin, J., Barna, N., Beckerly, R., Boguslavsky, L., Boukhalter, B., Brown, A., Castle, A., Colangelo, M., Collins, S., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Domino, M., Donnell, L., Doyle, M., Ferreira, P., Fitzhugh, W., Forrest, C., Funke, R., Gage, D., Galagan, J., Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L., Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Lehoczy, J., Liu, C., Locke, K., Macdonald, P., Margis, N., McEwan, P., McGurk, A., McKernan, K., McLaughlin, D., Meldrim, J., Morrow, J., Naylor, J., Norman, C. H., O'Connor, T., O'Donnell, P., Peterson, K., Pollara, V., Riley, R., Roy, A., Santos, R., Severy, P., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Tirrell, A., Vasilev, H., Vo, A., Wheeler, J., Wu, X., Wyman, D., Ye, W. J., Zimmer, A. and Zody, M.

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 Direct Submission  
 Submitted (17-NOV-1999) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 207661)  
 Birren, B., Nussbaum, C., Lander, E., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mhova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Riese, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vasilev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS  
 Direct Submission  
 Submitted (19-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA  
 4 (bases 1 to 207661)  
 Birren, B., Nussbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collymore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., Fitzgerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hafez, N., Hagos, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kamat, A., Karatas, A., Kelle, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J.,

TITLE  
 JOURNAL  
 COMMENT  
 All repeats were identified using RepeatMasker:  
 http://ftp.genome.washington.edu/RM/RepeatMasker.html  
 Smit, A.F.A. & Green, P. (1996-1997)  
 Center: Whitehead Institute/ MIT Center for Genome Research  
 Center code: MIBR  
 Web site: http://www-seq.wi.mit.edu  
 Contact: sequence\_submissions@genome.wi.mit.edu  
 Project Information  
 Center project name: I731  
 Center clone name: 216\_P\_6

FEATURES  
 source  
 Location/Qualifiers  
 1..207661  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /chromosome="17"  
 /map="17"  
 /clone="RP11-216P6"  
 /clone\_11b="RP11-11 Human Male BAC"  
 721  
 /note="Probably C"  
 2666..2923  
 /rpt\_family="AluXs"  
 3206..3412  
 /rpt\_family="AluUc"  
 3591..3888  
 /rpt\_family="AluXs"  
 complement(4652..4964)  
 /rpt\_family="AluXs"  
 5067..5109  
 /rpt\_family="Alu"  
 5106..5112  
 /rpt\_family="Alu"  
 /note="<30 qual SINGL region"  
 5110..5424  
 /rpt\_family="AluY"  
 5487..5582  
 /rpt\_family="MIR3"  
 complement(5846..5901)  
 /rpt\_family="MIR"  
 complement(5902..6209)  
 /rpt\_family="AluY"  
 complement(6210..6373)  
 /rpt\_family="MIR"  
 6847..7140  
 /rpt\_family="AluSg"  
 7860..7997  
 /rpt\_family="AluUc"  
 7998..8278  
 /rpt\_family="AluSg"  
 8279..8435  
 /rpt\_family="AluUc"  
 8436..8467  
 /rpt\_family="(CA)n"  
 9011..9302  
 /rpt\_family="AluSg"  
 9306..9597  
 /rpt\_family="AluXs"  
 complement(9849..10138)

Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mhova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Riese, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schupbach, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vasilev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

```

repeat_region /rpt_family="Aluub"
complement(10139, .10273)
/rpt_family="Aluub"
repeat_region complement(10285, .10579)
/rpt_family="Alu"
unSURE complement(10584, .10591)
/notes="30 qual SNL region"
repeat_region /rpt_family="AluSx"
10608, .10905
/rpt_family="AluSx"
repeat_region /rpt_family="AluSx"
11345, .11651
/rpt_family="AluSx"
repeat_region /rpt_family="Aluub"
11882, .12022
/rpt_family="Aluub"
repeat_region /rpt_family="AT_rich"
12040, .12072
/rpt_family="AT_rich"
repeat_region /rpt_family="CA)n"
12711, .12736
/rpt_family="CA)n"
repeat_region /rpt_family="AluSx"
12817, .13121
/rpt_family="AluSx"
repeat_region /rpt_family="AluY"
13292, .13590
/rpt_family="AluY"
repeat_region /rpt_family="L3"
14765, .15274
/rpt_family="L3"
repeat_region /rpt_family="AluYb8"
15275, .15322
/rpt_family="AluYb8"
repeat_region /rpt_family="AluSg"
15335, .15636
/rpt_family="AluSg"
repeat_region complement(15692, .15846)
/rpt_family="MER30"
repeat_region complement(15847, .16142)
/rpt_family="AluSp"
repeat_region complement(16143, .16218)
/rpt_family="MER30"
repeat_region /rpt_family="L3"
16220, .16432
/rpt_family="L3"
repeat_region /rpt_family="MLT1K"
16559, .16962
/rpt_family="MLT1K"
repeat_region /rpt_family="L1MC5"
17094, .17207
/rpt_family="L1MC5"
repeat_region /rpt_family="AluY"
17208, .17482
/rpt_family="AluY"
repeat_region /rpt_family="L1MC5"
17483, .17545
/rpt_family="L1MC5"
repeat_region /rpt_family="Aluub"
17905, .18082
/rpt_family="Aluub"
repeat_region /rpt_family="AluSx"
18152, .18462
/rpt_family="AluSx"
repeat_region complement(18647, .18889)
/rpt_family="MIR"
repeat_region complement(20440, .20741)
/rpt_family="AluSp"
repeat_region /rpt_family="AluY"
20925, .21221
/rpt_family="AluY"
repeat_region /rpt_family="AluSx"
21382, .21514
/rpt_family="AluSx"
repeat_region /rpt_family="AluSp"
21515, .21822
/rpt_family="AluSp"
repeat_region /rpt_family="AluSx"
21823, .21998
/rpt_family="AluSx"
repeat_region /rpt_family="AluY"
22246, .22543
/rpt_family="AluY"
repeat_region complement(23168, .23474)
/rpt_family="AluSx"

```

Query Match 1.7%; Score 52; DB 8; Length 207661;  
 Best Local Similarity 100.0%; Pred.No. 5.7e-16;  
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCGGATGATCACTGAGGCGAGGAGTTCAGAGCGAGCCCTGGCGACAT 2940  
 DB 116276 GAGGCGGATGATCACTGAGGCGAGGAGTTCAGAGCGAGCCCTGGCGACAT 116225

RESULT 193  
 AC027205/c  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

AC027205 211812 bp DNA linear HTG 24-AUG-2002  
 Homo sapiens chromosome 1 clone RP11-130F6 map 1, WORKING DRAFT  
 SEQUENCE, 51 unordered pieces.  
 AC027205  
 AC027205.2 GI:7652003  
 HTG: HTGS PHASE1: HTGS\_DRAFT.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homidae; Homo.

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS

1 (bases 1 to 211812)  
 Birren, B., Linton, L., Nusbaum, C. and Lander, E.  
 Homo sapiens chromosome 1, clone RP11-130F6  
 Unpublished  
 2 (bases 1 to 211812)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bede, F.,  
 Boguslavsky, L., Bouhgalter, B., Brown, A., Burkett, G.,  
 Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,  
 Collamore, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J.S.,  
 Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
 Grand-Pierre, N., Grant, G., Hagos, B., Jones, C., Kam, L., Karatas, A.,  
 Howard, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
 Klein, J., Lacroque, K., Lamazares, R., Lander, T., Lehoczy, J.,  
 Levine, R., Liu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,  
 McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McSheeters, R.,  
 Meldrum, J., Menus, L., Mihova, T., Miranda, C., Mleaga, V., Morrow, J.,  
 Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
 O'Sheil, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,  
 Plesni, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
 Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,  
 Vasilev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE  
 JOURNAL  
 REFERENCE  
 AUTHORS

Direct Submission  
 Submitted (28-MAR-2000) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 3 (bases 1 to 211812)  
 Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,  
 Anderson, S., Baldwin, J., Barna, N., Bastien, V., Bede, F.,  
 Boguslavsky, L., Bouhgalter, B., Brown, A., Burkett, G.,  
 Campopiano, A., Castle, A., Choepel, Y., Colangelo, M., Collins, S.,  
 Collamore, A., Cooke, P., Dearrellano, K., Dewar, K., Diaz, J.S.,  
 Dodge, S., Domingo, M., Doyle, M., Ferreira, P., Fitzhugh, W., Gage, D.,  
 Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,  
 Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L.,  
 Howard, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,  
 Klein, J., Lacroque, K., Lamazares, R., Lander, T., Lehoczy, J.,  
 Levine, R., Liu, C., Liu, G., Locke, K., MacDonald, P., Margulis, N.,  
 McCarthy, M., McEwan, P., McGuirk, A., McKernan, K., McSheeters, R.,  
 Meldrum, J., Menus, L., Mihova, T., Miranda, C., Mleaga, V., Morrow, J.,  
 Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,  
 O'Sheil, D., Oliver, T.M., Oliver, J., Peterson, K., Pierre, N.,  
 Plesni, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,  
 Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,  
 Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,  
 Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,  
 Vasilev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,  
 Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission  
 Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome  
 Research, 320 Charles Street, Cambridge, MA 02141, USA  
 On Apr 27, 2000 this sequence version replaced gi:7331532.  
 All repeats were identified using RepeatMasker:  
 smc, A.F.A. & Green, P. (1996-1997)  
 http://ftp.genome.washington.edu/RW/RepeatMasker.html  
 ----- Genome Center  
 Center: Whitehead Institute/ MIT Center for Genome Research

```
Center code: W1BR
Web site: http://www-seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information -----
Center project name: L7762
Center clone name: 130_P_6
----- Summary Statistics -----
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 183012 bases at least Q40
Consensus quality: 197888 bases at least Q30
Consensus quality: 203091 bases at least Q20
Insert size: 206812; sum-of-ctngs
Quality coverage: 3.6 in Q20 bases; sum-of-ctngs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 51 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
1012: contig of 1012 bp in length
1112: gap of 100 bp
1113: contig of 1249 bp in length
2362: gap of 100 bp
2461: contig of 1204 bp in length
3665: gap of 100 bp
3765: contig of 1248 bp in length
5013: gap of 100 bp
5113: contig of 1572 bp in length
5114: gap of 100 bp
6685: gap of 100 bp
6785: contig of 1203 bp in length
6886: gap of 100 bp
7988: contig of 1423 bp in length
8088: gap of 100 bp
8089: contig of 1273 bp in length
9512: gap of 100 bp
9513: contig of 1238 bp in length
10884: gap of 100 bp
10885: contig of 1232 bp in length
10985: gap of 100 bp
12322: contig of 1113 bp in length
12323: gap of 100 bp
13435: contig of 1258 bp in length
13436: gap of 100 bp
14793: contig of 1181 bp in length
14794: gap of 100 bp
14894: contig of 1168 bp in length
16074: gap of 100 bp
16075: contig of 1069 bp in length
16174: gap of 100 bp
17342: contig of 1540 bp in length
17343: gap of 100 bp
17442: contig of 1350 bp in length
17443: gap of 100 bp
18511: contig of 1935 bp in length
18512: gap of 100 bp
18612: contig of 1669 bp in length
20151: gap of 100 bp
20152: contig of 1350 bp in length
20252: gap of 100 bp
21601: contig of 1249 bp in length
21602: gap of 100 bp
23636: contig of 1559 bp in length
23637: gap of 100 bp
23736: contig of 1659 bp in length
25405: gap of 100 bp
25406: contig of 1148 bp in length
25505: gap of 100 bp
26653: contig of 1249 bp in length
26753: gap of 100 bp
26754: contig of 1575 bp in length
28002: gap of 100 bp
28102: contig of 1575 bp in length
28103: gap of 100 bp
29661: contig of 1259 bp in length
29662: gap of 100 bp
31020: contig of 1638 bp in length
31021: gap of 100 bp
32758: contig of 1575 bp in length
32759: gap of 100 bp
32858: contig of 1575 bp in length
34433: gap of 100 bp
34434: contig of 2266 bp in length
36799: gap of 100 bp
```

```
36800 36899: gap of 100 bp
36900 38788: contig of 1889 bp in length
38789 38888: gap of 100 bp
38889 41490: contig of 2602 bp in length
41491 41590: gap of 100 bp
41591 43768: contig of 2178 bp in length
43769 43868: gap of 100 bp
43869 45634: contig of 1766 bp in length
45635 45734: gap of 100 bp
45735 46330: contig of 556 bp in length
46331 46430: gap of 100 bp
46431 48111: contig of 1681 bp in length
48112 48211: gap of 100 bp
48212 50798: contig of 2587 bp in length
50799 50898: gap of 100 bp
50899 51541: contig of 643 bp in length
51542 51641: gap of 100 bp
51642 54229: contig of 2588 bp in length
54230 54329: gap of 100 bp
54330 57592: contig of 3263 bp in length
57593 57693: gap of 100 bp
57693 60380: contig of 2688 bp in length
60381 60480: gap of 100 bp
60481 64861: contig of 4381 bp in length
64862 64961: gap of 100 bp
64961 67169: contig of 2208 bp in length
67170 67269: gap of 100 bp
67270 70318: contig of 3049 bp in length
70319 70418: gap of 100 bp
70419 74740: contig of 4332 bp in length
74741 74840: gap of 100 bp
74841 79050: contig of 4210 bp in length
79051 79150: gap of 100 bp
79151 84593: contig of 5443 bp in length
84594 84693: gap of 100 bp
84694 94483: contig of 9790 bp in length
94484 94584: gap of 100 bp
94584 105653: contig of 11070 bp in length
105654 105753: gap of 100 bp
105754 116723: contig of 10970 bp in length
116724 116824: gap of 100 bp
116824 129473: contig of 12650 bp in length
129474 129573: gap of 100 bp
129574 146634: contig of 17061 bp in length
146635 146734: gap of 100 bp
146735 162593: contig of 1559 bp in length
162594 162793: gap of 100 bp
162794 178165: contig of 15372 bp in length
178166 178265: gap of 100 bp
178266 196145: contig of 17880 bp in length
196146 196245: gap of 100 bp
196246 211812: contig of 15567 bp in length.

FEATURES
    source
        1..211812
            /organism="Homo sapiens"
            /mol_type="genomic DNA"

Query Match 1.7%; Score 52; DB 14; Length 211812;
Best Local Similarity 100.0%; Fred. No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGCGCACTGCACTCCAGGCTGGGCAAGAGCAAGACTGTCTC 3122
Db 211005 CAAGATTGCGCACTGCACTCCAGGCTGGGCAAGAGCAAGACTGTCTC 210954

RESULT 194
AC135988/c 213947 bp DNA linear RTG 21-NOV-2002
LOCUS Homo sapiens chromosome 15 clone RP11-701021 map 15, WORKING DRAFT
DEFINITION SEQUENCE, 4 unordered pieces.
ACCESSION AC135988
VERSION AC135988.2 GI:25141060
```

```
KEYWORDS      HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FUILLTOP.
SOURCE         Homo sapiens (human)
ORGANISM       Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
               Homiidae; Homo.
REFERENCE      1 (bases 1 to 213947)
AUTHORS        Birren,B., Nuebaum,C. and Lander,E.
TITLE          Homo sapiens chromosome 15, clone RP11-701021
JOURNAL        Unpublished
REFERENCE      2 (bases 1 to 213947)
AUTHORS        Birren,B., Nuebaum,C., Lander,E., All,A., Allen,N., Anderson,S.,
               Barre,N., Baetien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,
               Camarata,U., Chang,U., Chazaro,B., Choepel,Y., Collymore,A.,
               Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
               Gadya,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
               Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
               Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
               Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J.,
               Matthews,C., McCarthy,M., Meldrim,J., Menue,L., Mihova,T.,
               Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
               Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
               Peterson,K., Phunhphang,P., Pierre,N., Raymond,C., Retta,R.,
               Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
               Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
               Stojanovic,N., Talamas,J., Testfaye,S., Theodore,J., Topham,K.,
               Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
               Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
JOURNAL        Submitted (26-OCT-2002) Whitehead Institute/MIT Center for Genome
REFERENCE      Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS        3 (bases 1 to 213947)
               Birren,B., Nuebaum,C., Lander,E., All,A., Allen,N., Anderson,S.,
               Barre,N., Baetien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,
               Camarata,U., Chang,U., Chazaro,B., Choepel,Y., Collymore,A.,
               Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
               Gadya,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,
               Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
               Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
               Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J.,
               Matthews,C., McCarthy,M., Meldrim,J., Menue,L., Mihova,T.,
               Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C.,
               Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,
               Peterson,K., Phunhphang,P., Pierre,N., Raymond,C., Retta,R.,
               Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupback,R.,
               Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N.,
               Stojanovic,N., Talamas,J., Testfaye,S., Theodore,J., Topham,K.,
               Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,
               Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
DIRECT SUBMISSION
JOURNAL        Submitted (21-NOV-2002) Whitehead Institute/MIT Center for Genome
REFERENCE      Research, 320 Charles Street, Cambridge, MA 02141, USA
AUTHORS        On Nov 21, 2002 this sequence version replaced gi:24414513.
COMMENT        All repeats were identified using RepeatMasker:
               Smit, A.F.A. & Green, P. (1996-1997)
               http://ftp.genome.washington.edu/RM/RepeatMasker.html
               ----- Genome Center
               Center: Whitehead Institute/ MIT Center for Genome Research
               Web site: http://www-seg.wi.mit.edu
               Center code: WIBR
               Contact: sequence_submissions@genome.wi.mit.edu
               ----- Project Information
               Center project name: 701_O_21
               L28438
               ----- Summary Statistics
               Sequencing vector: Plasmid; n/a; 100% of reads
               Chemistry: Dye-terminator Big Dye; 100% of reads
               Assembly program: Phrap; version 0.960731
               Consensus quality: 213175 bases at least Q40
               Consensus quality: 213476 bases at least Q30
               Consensus quality: 213580 bases at least Q20
```

REFERENCE  
AUTHORS  
2 (bases 1 to 215780)  
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,  
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,  
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Colliamore,A.,  
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,  
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,  
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
Kamit,A., Karatas,A., Kells,C., Landers,T., Levine,R.,  
Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J.,  
Mathews,C., McCarthy,M., Meldrum,J., Menus,L., Mihova,T.,  
Mlenga,V., Murphy,T., Nguyen,C., Nicol,R., Norbu,C.,  
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,  
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,  
Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupbach,R.,  
Seaman,S., Severy,P., Smith,C., Spencer,B., Strange-Thomann,N.,  
Stojanovic,N., Talamas,J., Testaye,S., Theodore,J., Topham,K.,  
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,  
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE  
JOURNAL  
Submitted (26-OCT-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
3 (bases 1 to 215780)

REFERENCE  
AUTHORS  
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,  
Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B.,  
Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Colliamore,A.,  
Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,  
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,  
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hafez,N.,  
Hagos,B., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,  
Kamit,A., Karatas,A., Kells,C., Landers,T., Levine,R.,  
Lindblad-Toh,K., Liu,G., Maclean,C., Macdonald,P., Major,J.,  
Mathews,C., McCarthy,M., Meldrum,J., Menus,L., Mihova,T.,  
Mlenga,V., Murphy,T., Nguyen,C., Nicol,R., Norbu,C.,  
Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J.,  
Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R.,  
Rise,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupbach,R.,  
Seaman,S., Severy,P., Smith,C., Spencer,B., Strange-Thomann,N.,  
Stojanovic,N., Talamas,J., Testaye,S., Theodore,J., Topham,K.,  
Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X.,  
Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.

TITLE  
JOURNAL  
Submitted (21-NOV-2002) Whitehead Institute/MIT Center for Genome  
Research, 320 Charles Street, Cambridge, MA 02141, USA  
On Nov 21, 2002 this sequence version replaced gi:2414514.  
All repeats were identified using RepeatMasker:  
Smit,A.F.A. & Green, P. (1996-1997)  
http://ftp.genome.washington.edu/RM/RepeatMasker.html

COMMENT  
----- Genome Center  
Center: Whitehead Institute/ MIT Center for Genome Research  
Center code: WIBR  
Web site: http://www-seq.wi.mit.edu  
Contact: sequence\_submissions@genome.wi.mit.edu  
----- Project Information  
Center project name: L28440  
Center clone name: 800 O 12  
----- Summary Statistics  
Sequencing vector: Plasmid; n/a; 100% of reads  
Chemistry: Dye-terminator Big Dye; 100% of reads  
Assembly program: Phrap; version 0.960731  
Consensus quality: 215088 bases at least Q40  
Consensus quality: 215264 bases at least Q30  
Consensus quality: 215336 bases at least Q20  
Insert size: 215000; agarose-efp  
Insert size: 215480; sum-of-ctnigs  
Quality coverage: 15.5 in Q20 bases; agarose-efp  
Quality coverage: 15.5 in Q20 bases; sum-of-ctnigs  
-----  
\* NOTE: This is a 'working draft' sequence. It currently  
\* consists of 4 contigs. The true order of the pieces  
\* is not known and their order in this sequence record is  
\* arbitrary. Gaps between the contigs are represented as  
\* runs of N, but the exact sizes of the gaps are unknown.

\* This record will be updated with the finished sequence  
\* as soon as it is available and the accession number will  
\* be preserved.  
\* 1 60439: contig of 60439 bp in length  
\* 60440 60539: gap of 100 bp  
\* 60540 61230: contig of 691 bp in length  
\* 61231 61330: gap of 100 bp  
\* 61331 61917: contig of 587 bp in length  
\* 61918 62017: gap of 100 bp  
\* 62018 215780: contig of 153763 bp in length.  
Location/Qualifiers  
1..215780  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="15"  
/map="15"  
/clone="RP11-800012"  
/clone\_id="RP11-800012 Human Male BAC"  
1..60439  
/note="assembly\_fragment"  
clone\_end:86  
vector\_side:left"  
60440..60539  
/estimated\_length=100  
60540..61230  
/note="assembly\_fragment"  
61231..61330  
/estimated\_length=100  
61331..61917  
/note="assembly\_fragment"  
61918..62017  
/estimated\_length=100  
62018..215780  
/note="assembly\_fragment"

ORIGIN  
Query Match 1.7%; Score 52; DB 14; Length 215780;  
Best Local Similarity 100.0%; Pred. No. 5.7e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Qy 2889 GAGCAGGTGATCACTGAGGCGCAGGATTGAGACCGAGCTGCGCAACAT 2940  
Db 6909 GAGCAGGTGATCACTGAGGCGCAGGATTGAGACCGAGCTGCGCAACAT 6960  
|||||  
RESULT 196  
AC002429/c 234053 bp DNA linear PRI 03-OCT-2003  
LOCUS  
DEFINITION  
AC002429 Homo sapiens BAC clone GSI-200K5 from 7, complete sequence.  
AC002429.1 GI:2335067  
VERSION  
KEYWORDS  
HTG.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominoidea; Homo.  
1 (bases 1 to 234053)  
Sultson,J.E. and Wilson,R.  
Toward a complete human genome sequence  
JOURNAL  
Genome Res. 8 (11), 1097-1108 (1998)  
PUBMED  
9847074  
2 (bases 1 to 234053)  
Duckele,G., Graves,T. and Hawkins,M.  
The sequence of Homo sapiens BAC clone GSI-200K5  
JOURNAL  
Unpublished (2001)  
REFERENCE  
3 (bases 1 to 234053)  
Waterston,R.  
TITILE  
JOURNAL  
Submitted (19-AUG-1997) Department of Genetics, Washington  
University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
REFERENCE  
4 (bases 1 to 234053)



AUTHORS Waterston,R.  
TITLE Direct Submission  
JOURNAL Submitted (03-FEB-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
REFERENCE 5 (bases 1 to 234053)  
AUTHORS Waterston,R.  
TITLE Direct Submission  
JOURNAL Submitted (04-FEB-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
REFERENCE 6 (bases 1 to 234053)  
AUTHORS Waterston,R.  
TITLE Direct Submission  
JOURNAL Submitted (26-APR-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
REFERENCE 7 (bases 1 to 234053)  
AUTHORS Wilson,R.  
TITLE Direct Submission  
JOURNAL Submitted (03-OCT-2003) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA  
COMMENT ----- Genome Center  
Center: Washington University Genome Sequencing Center  
Center code: WUGSC  
Web site: <http://genome.wustl.edu>  
Contact: saplens@wuston.wustl.edu  
----- Summary Statistics  
Center project name: H\_GS200K05  
-----  
NOTICE: This sequence may not represent the entire insert of this clone. It may be shorter because we only sequence overlapping clone sections once, or longer because we provide a small overlap between neighboring data submissions.  
  
This sequence was finished as follows unless otherwise noted:  
all regions were double stranded, sequenced with an alternate chemistry, or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one subclone; and the assembly was confirmed by restriction digest.  
  
MAPPING INFORMATION:  
The sequence of this clone was established as part of a mapping and sequencing collaboration between the NHGRI Chromosome 7 Mapping Project (Eric D. Green, Director), John D. McPherson in the Department of Genetics (Washington University), and the Washington University Genome Sequencing Center. For additional information about the map position of this sequence, see <http://www.nhgri.nih.gov/DIR/GRB/CHR7>, send <mailto:egreen@nhgri.nih.gov>, or see <http://genome.wustl.edu>  
  
SOURCE INFORMATION:  
This clone is from the first BAC library from Genome Systems, Inc. (<http://www.genomesystems.com>).  
Cell line: lymphoblastoid  
Haplotypes: two  
VECTOR: pBelBAC  
Selection: chloramphenicol.  
Location/Qualifiers  
1. 234053  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/chromosome="7"  
/map="7"  
/clone="GS1-200K5"  
/clone\_1lb="GSBAC1"  
/map\_1lb="81"  
repeat\_region /rpt\_family="CT-rich"  
477..628  
repeat\_region /rpt\_family="MIR"  
2145..2168  
repeat\_region /rpt\_family="AT-rich"

repeat\_region 3013..3036  
/rpt\_family="AT-rich"  
repeat\_region 3230..3538  
/rpt\_family="L1"  
repeat\_region 3539..3625  
/rpt\_family="MER2\_type"  
repeat\_region 3636..4018  
/rpt\_family="MER2\_type"  
repeat\_region 4040..4191  
/rpt\_family="L1"  
repeat\_region 7520..7597  
/rpt\_family="Alu"  
repeat\_region 7598..8047  
/rpt\_family="L1"  
repeat\_region 8048..8335  
/rpt\_family="Alu"  
misc\_feature 8073..8336  
/note="CpG island (%GC=60.6, o/e=0.88, #CpGs=21)"  
repeat\_region 8336..8580  
/rpt\_family="L1"  
repeat\_region 10511..10654  
/rpt\_family="Alu"  
repeat\_region 10869..10970  
/rpt\_family="L2"  
repeat\_region 13585..13894  
/rpt\_family="Alu"  
repeat\_region 14043..14102  
/rpt\_family="MER1\_type"  
repeat\_region 15225..15261  
/rpt\_family="(T)n"  
repeat\_region 16103..16139  
/rpt\_family="AT-rich"  
repeat\_region 16215..16474  
/rpt\_family="Alu"  
repeat\_region 16669..16826  
/rpt\_family="L2"  
repeat\_region 16838..17493  
/rpt\_family="L2"  
repeat\_region 17533..17630  
/rpt\_family="L2"  
repeat\_region 17631..17938  
/rpt\_family="Alu"  
repeat\_region 17939..18058  
/rpt\_family="L2"  
repeat\_region 18676..18974  
/rpt\_family="Alu"  
repeat\_region 19969..20188  
/rpt\_family="MER1\_type"  
repeat\_region 20255..20328  
/rpt\_family="MER1\_type"  
repeat\_region 20349..20503  
/rpt\_family="Alu"  
repeat\_region 20506..20572  
/rpt\_family="MER1\_type"  
repeat\_region 20764..20907  
/rpt\_family="L2"  
repeat\_region 21114..21414  
/rpt\_family="Alu"  
repeat\_region 25771..25861  
/rpt\_family="CRI"  
repeat\_region 26146..26170  
/rpt\_family="AT-rich"  
repeat\_region 26220..26642  
/rpt\_family="L1"  
repeat\_region 27343..27555  
/rpt\_family="L1"  
repeat\_region 28645..28672  
/rpt\_family="(CATATA)n"  
repeat\_region 28916..29059  
/rpt\_family="MER1\_type"  
repeat\_region 29080..29472  
/rpt\_family="ERV1"  
repeat\_region 29524..29577

```
repeat_region      /rpt_family="MER1_type"
30140 .30236
repeat_region      /rpt_family="MIR"
30312 .31450
repeat_region      /rpt_family="L1"
31525 .31829
repeat_region      /rpt_family="Alu"
31873 .32741
repeat_region      /rpt_family="L1"
32761 .33823
repeat_region      /rpt_family="L1"
33825 .34218
repeat_region      /rpt_family="ERV1"
34230 .34289
repeat_region      /rpt_family="ERV1"
34304 .34494
repeat_region      /rpt_family="L1"
34495 .34789
repeat_region      /rpt_family="Alu"
34790 .35979
repeat_region      /rpt_family="L1"
35977 .36335
repeat_region      /rpt_family="L1"
36342 .36677
repeat_region      /rpt_family="L1"
36678 .36981
repeat_region      /rpt_family="Alu"
36982 .37759
repeat_region      /rpt_family="L1"
```

```
Query Match      1.7%; Score 52; DB 8; Length 234053;
Best Local Similarity 100.0%; Pred.No. 5.7e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      3071 CAAGATTGTCACCTGCTCCAGCTGCGCAACAGAGCAAGACTGTCTC 3122
Db      223398 CAAGATTGTCACCTGCTCCAGCTGCGCAACAGAGCAAGACTGTCTC 223347
```

```
RESULT 197
AC145896      256000 bp      DNA      linear      HTG 13-AUG-2003
LOCUS      Pan troglodytes chromosome UNK clone RP43-8H23, *** SEQUENCING IN
DEFINITION
AC145896
AC145896.1 GI:33386828
VERSION
AC145896.1 GI:33386828
KEYWORDS
HTG; HTGS_PHASE1.
SOURCE
Pan troglodytes (chimpanzee)
ORGANISM
Pan troglodytes
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominoidea; Pan.
1 (bases 1 to 256000)
REFERENCE
AUTHORS      Wilson,R.K.
TITLE      The sequence of Pan troglodytes clone
JOURNAL      Unpublished
REFERENCE
AUTHORS      Wilson,R.K.
TITLE      Direct Submision
JOURNAL      Submitted (01-AUG-2003) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA
3 (bases 1 to 256000)
REFERENCE
AUTHORS      Wilson,R.K.
TITLE      Direct Submision
JOURNAL      Submitted (13-AUG-2003) Genetics, Genome Sequencing Center, 4444
Forest Park Parkway, St. Louis, MO 63108, USA
```

```
COMMENT
----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: C_PT008H23
```

```
----- Summary Statistics -----
Sequencing vector: M13; 0%
Sequencing vector: plasmid; 100%
Chemistry: Dye-primer ET; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 237274 bases at least Q40
Consensus quality: 240505 bases at least Q30
Consensus quality: 242221 bases at least Q20
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 48 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1      1185: contig of 1185 bp in length
*
*      1186: gap of unknown length
*
*      1286: gap of unknown length
*
*      1286: contig of 1184 bp in length
*
*      2470: gap of unknown length
*
*      2570: contig of 1410 bp in length
*
*      3980: gap of unknown length
*
*      4080: contig of 1216 bp in length
*
*      5296: gap of unknown length
*
*      5396: gap of unknown length
*
*      6631: gap of unknown length
*
*      6730: gap of unknown length
*
*      7921: contig of 1190 bp in length
*
*      8020: gap of unknown length
*
*      8921: contig of 1432 bp in length
*
*      9453: gap of unknown length
*
*      9553: contig of 1253 bp in length
*
*      10806: gap of unknown length
*
*      10806: contig of 1240 bp in length
*
*      12146: gap of unknown length
*
*      12245: gap of unknown length
*
*      13423: contig of 1178 bp in length
*
*      13424: gap of unknown length
*
*      13524: contig of 1503 bp in length
*
*      15026: gap of unknown length
*
*      15027: gap of unknown length
*
*      15126: gap of unknown length
*
*      16408: contig of 1282 bp in length
*
*      16508: gap of unknown length
*
*      16509: contig of 1392 bp in length
*
*      17901: gap of unknown length
*
*      18000: gap of unknown length
*
*      18001: contig of 1931 bp in length
*
*      19931: gap of unknown length
*
*      19932: gap of unknown length
*
*      20032: contig of 1188 bp in length
*
*      21220: gap of unknown length
*
*      21319: gap of unknown length
*
*      21320: contig of 2111 bp in length
*
*      23430: gap of unknown length
*
*      23431: contig of 1544 bp in length
*
*      23531: gap of unknown length
*
*      25074: contig of 1544 bp in length
*
*      25174: gap of unknown length
*
*      25175: gap of unknown length
*
*      26544: contig of 1370 bp in length
*
*      26545: gap of unknown length
*
*      28429: gap of unknown length
*
*      28429: contig of 1765 bp in length
*
*      28529: gap of unknown length
*
*      28530: gap of unknown length
*
*      28530: contig of 1380 bp in length
*
*      29910: gap of unknown length
*
*      30009: gap of unknown length
*
*      31609: contig of 1600 bp in length
*
*      31610: gap of unknown length
*
*      31710: gap of unknown length
*
*      34227: contig of 2517 bp in length
*
*      34326: gap of unknown length
*
*      34327: gap of unknown length
*
*      35499: contig of 1173 bp in length
*
*      35599: gap of unknown length
*
*      35600: gap of unknown length
*
*      36828: contig of 1229 bp in length
*
*      36929: gap of unknown length
*
*      39530: contig of 2602 bp in length
*
*      39531: gap of unknown length
*
*      39631: gap of unknown length
*
*      40715: contig of 1065 bp in length
*
*      40815: gap of unknown length
*
*      42410: contig of 1595 bp in length
*
*      42411: gap of unknown length
```

	*	42511	43677:	contig of 1167 bp in length
	*	43678	43777:	gap of unknown length
	*	43778	44933:	contig of 1156 bp in length
	*	44933	45033:	gap of unknown length
	*	45034	45566:	contig of 1533 bp in length
	*	46567	46666:	gap of unknown length
	*	46667	48332:	contig of 1666 bp in length
	*	48333	48432:	gap of unknown length
	*	48433	50446:	contig of 2014 bp in length
	*	50447	50546:	gap of unknown length
	*	50547	51956:	contig of 1410 bp in length
	*	51957	52056:	gap of unknown length
	*	52057	54338:	contig of 2282 bp in length
	*	54339	54438:	gap of unknown length
	*	54439	56408:	contig of 1970 bp in length
	*	56409	56508:	gap of unknown length
	*	60180	60179:	contig of 3671 bp in length
	*	60280	60279:	gap of unknown length
	*	64270	64269:	contig of 3990 bp in length
	*	64370	64369:	gap of unknown length
	*	65820	65819:	contig of 1450 bp in length
	*	65920	65919:	gap of unknown length
	*	70604	70603:	contig of 4684 bp in length
	*	70704	74447:	gap of unknown length
	*	74448	74447:	contig of 3744 bp in length
	*	74548	80301:	gap of unknown length
	*	80302	80301:	contig of 5754 bp in length
	*	80402	93561:	gap of unknown length
	*	93562	93561:	contig of 1160 bp in length
	*	93662	93661:	gap of unknown length
	*	107828	107828:	contig of 14167 bp in length
	*	107929	129790:	gap of unknown length
	*	129691	129790:	contig of 21762 bp in length
	*	129791	150182:	gap of unknown length
	*	150183	150282:	contig of 20392 bp in length
	*	150283	181315:	gap of unknown length
	*	181316	181415:	gap of unknown length
	*	207034	207034:	contig of 25619 bp in length
	*	207035	207134:	gap of unknown length
	*	207135	256000:	contig of 48866 bp in length
FEATURES			Location/Qualifiers	
source			1..256000	
			/organism="Pan troglodytes"	
			/mol_type="genomic DNA"	
			/db_xref="taxon:9598"	
			/chromosome="20K"	
			/clone="RP43-BH23"	
			1..1185	
misc_feature			/note="assembly_name:Contig23"	
gap			1186..1285	
			/estimated_length=unknown	
misc_feature			1286..2463	
			/note="assembly_name:Contig25"	
gap			2470..2569	
			/estimated_length=unknown	
misc_feature			2570..3975	
			/note="assembly_name:Contig31"	
gap			3980..4079	
			/estimated_length=unknown	
misc_feature			4080..5295	
			/note="assembly_name:Contig33"	
gap			5296..5395	
			/estimated_length=unknown	
misc_feature			5396..6630	
			/note="assembly_name:Contig37"	
gap			6631..6730	
			/estimated_length=unknown	
misc_feature			6731..7920	
			/note="assembly_name:Contig38"	
gap			7921..8020	
			/estimated_length=unknown	
misc_feature			8021..9452	

[illegible]

FEATURES	
source	location/Qualifiers
gap	1..279430
gap	/organism="Pan troglodytes"
gap	/mol_type="genomic DNA"
gap	/db_xref="taxon:9598"
gap	/clone="rp43-47k21"
gap	/clove_lib="RPC1 - 43 Male Chimpanzee BAC library"
gap	2055..2154
gap	/estimated_length=unknown
gap	5486..5585
gap	/estimated_length=unknown
gap	8524..8623
gap	/estimated_length=unknown
gap	11067..11166
gap	/estimated_length=unknown
gap	14581..14680
gap	/estimated_length=unknown
gap	17964..18063
gap	/estimated_length=unknown
gap	21382..21481
gap	/estimated_length=unknown
gap	25968..26067
gap	/estimated_length=unknown
gap	30351..30450
gap	/estimated_length=unknown
gap	37059..37158
gap	/estimated_length=unknown
gap	44121..44220
gap	/estimated_length=unknown
gap	56125..56224
gap	/estimated_length=unknown
gap	70496..70595
gap	/estimated_length=unknown
gap	88933..89031

Db 273885 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 273936

RESULT 200  
 BV634964/C  
 LOCUS  
 DEFINITION  
 BV634964 715 bp DNA linear STS 15-APR-2005  
 S21P6621FB7.T0 Noemie Pan troglodytes troglodytes STS genomic,  
 sequence tagged site.  
 ACCESSION  
 BV634964  
 VERSION  
 BV634964.1 GI:62625142  
 KEYWORDS  
 STS.  
 SOURCE  
 ORGANISM  
 Pan troglodytes troglodytes  
 Pan troglodytes troglodytes  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homiidae; Pan.  
 REFERENCE  
 1 (bases 1 to 715)  
 Mikkelsen,T.S., Hillier,W.L., Eichler,E.E., Zody,M.C. and  
 Jaffe,D.B.  
 TITLE  
 Initial Sequence of the Chimpanzee Genome and Comparison with the  
 Human Genome  
 JOURNAL  
 Unpublished (2005)  
 COMMENT  
 Contact: Michael C. Zody  
 Broad Institute of MIT and Harvard  
 320 Charles Street, Cambridge, MA 02141, USA  
 Tel: 6172580933  
 Fax: 6172580903  
 Email: mczody@broad.mit.edu  
 Primer A: No sequence submitted  
 Primer B: No sequence submitted  
 STS size: 715  
 Protocol:  
 23,021,928 chimpanzee whole genome shotgun reads were aligned to  
 the Human genome NCBI  
 Build 34 (hg16, July 2003). Chimp WGS reads were from 9 donors,  
 including Clint (Pan  
 troglodytes verus), 3 other Pan troglodytes verus chimps  
 (Donald,Karlen,Yvonne), 3 Pan  
 troglodytes troglodytes chimps (Noemie,Masuku,Clara) and 2 chimps  
 of unknown origin  
 (Gon,Unknown Chimp). Common names: Pan troglodytes verus is the  
 western chimp and Pan  
 troglodytes troglodytes is the central chimp. To be included in  
 chimpanzee SNP discovery, a  
 read must be at least 500bp in length, at least 50% of its base  
 calls must have Phred  
 score >= 20, at least 30% of its base calls must satisfy  
 SNQs(30,25) (single strand NQS, the  
 base in question has Phred score >= 30, the surrounding 10 bases in  
 the read have Phred  
 score >= 25), and the read must have at least 200 bp SNQs(30,25)  
 bases. Reads not uniquely  
 placed in the genome and read pairs whose two ends were not  
 consistently placed were  
 discarded. After above filtering, NQS(30,25) standard was applied  
 to all pairs of  
 overlapping reads to call NQS bases and SNPs. Alignments (between  
 two reads) with less  
 than 100 NQS bases or with SNP rate > 0.01 were discarded. To  
 exclude alignment between two  
 copies of a single read, comparisons between two reads that share  
 95% of their genome  
 alignments (>=95% bases of read A and >=95% bases of read B were  
 placed at the same locus  
 of human genome) were discarded.  
 Location/Qualifiers  
 1..715  
 /organism="Pan troglodytes troglodytes"  
 /mol\_type="genomic DNA"  
 /sub\_species="troglodytes"  
 /db\_xref="taxon:37011"  
 /clone\_lib="Noemie"  
 <1..>715

Query Match 1.6%; Score 51; DB 10; Length 715;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-15;  
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 3072 AAGATTGCGCACTGCCTCCAGCCTGGGCAAGAGCAAGACTGTCTC 3122  
 |||||  
 Db 480 AAGATTGCGCACTGCCTCCAGCCTGGGCAAGAGCAAGACTGTCTC 430

Search completed: May 11, 2006, 08:43:47  
 Job time : 15448 secs

# FEATURES source

STS  
 ORIGIN

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2006, 04:27:18 ; Search time 11817 Seconds  
(without alignments)  
12360.944 Million cell updates/sec

Title: US-10-760-320A-102

Perfect score: 3122

Sequence: 1 actagagctgsggttagcgc.....acagagcaagactctgtctc 3122

Scoring table: OLIGO\_NUC

Gapop 60.0 , Gapext 60.0

Searched: 41078325 seqs, 2339354128 residues

Word size: 1

Total number of hits satisfying chosen parameters: 82156240

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 500 summaries

Database: EST:\*

1: gb\_est1:\*\n2: gb\_est2:\*\n3: gb\_est3:\*\n4: gb\_est4:\*\n5: gb\_est5:\*\n6: gb\_est6:\*\n7: gb\_est7:\*\n8: gb\_est8:\*\n9: gb\_est9:\*\n10: gb\_est10:\*\n11: gb\_est11:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	646	20.7	660	5	BX116028
2	613	19.6	670	6	CF145408 UI-HF-CH0
3	598	19.2	677	6	CF145448 UI-HF-CH0
4	531	17.0	685	3	BM671616 UI-E-CO1-
5	519	16.6	578	3	BM707056 UI-E-CR1-
6	500	16.0	500	3	BM710194 UI-E-CO1-
7	450	14.4	655	3	BQ187216 UI-E-EJ1-
8	427	13.7	476	1	AM302149 x801E06.x
9	413	13.2	455	1	AI375213 TC10E06.x
10	412	13.2	480	3	BM673230 UI-E-CR1-
11	395	12.7	972	5	BO682843 AGENCOURT
12	348	11.1	1376	3	BM622937 AGENCOURT
13	333	10.7	595	8	W81663 zdb8e06.t1
14	290	9.3	728	3	BI458731 60319879
15	180	5.8	605	5	BX645932 DKFZp781B
16	144	4.6	449	8	W81222 zdb8e06.t1
17	53	1.7	709	10	AG013685 Homo sapi
18	53	1.7	718	10	AG013687 Homo sapi
19	53	1.7	718	10	AG013686 Homo sapi
20	52	1.7	354	1	AA679936 ac85g01.s
21	52	1.7	441	1	AI732120 ac85g01.x
22	52	1.7	441	1	AI732180 ac85g01.x



96	45	726	10	AG009127	Homo sapi	159	44	1.4	637	5	BX482192	BX482192	DKFZP686H
97	45	731	10	AG009138	Homo sapi	170	44	1.4	638		AG097816	AG097816	HS_5536_B
98	45	741	10	AG009129	Homo sapi	171	44	1.4	641	1	AV733586	AV733586	AV733586
99	45	754	5	BX457023	BX457023	172	44	1.4	649	10	AG0255905	AG0255905	Homo sapi
100	45	768	2	BG260049	602371633	173	44	1.4	650	10	AG080818	AG080818	Pan trogl
101	45	781	5	BX411269	BX411269	174	44	1.4	652	1	AV722027	AV722027	AV722027
102	45	830	6	CD242479	AGENCOURT	175	44	1.4	654	10	AG106557	AG106557	Pan trogl
103	45	856	5	BU959380	AGENCOURT	176	44	1.4	655	1	AM959999	AM959999	ESR172070
104	45	874	5	BZ771376	mcg76c11.	177	44	1.4	694	9	AO740363	AO740363	HS_5501_A
105	45	948	6	CF596843	AGENCOURT	178	44	1.4	706	5	AO616112	AO616112	UI-H-DT0-
106	45	988	5	BO896885	AGENCOURT	179	44	1.4	696	5	AO626202	AO626202	CITB1-S1-
107	45	1123	5	BX404721	BX404721	180	44	1.4	715	5	BX504393	BX504393	DKFZP686O
108	45	203	2	BG959135	PM4-CT080	181	44	1.4	716	5	BX098422	BX098422	BX098422
109	44	219	2	BG010132	MR3-GN018	182	44	1.4	745	10	AG175100	AG175100	Pan trogl
110	44	281	2	BF893386	CV3-MT012	183	44	1.4	781	6	CB308313	CB308313	AGENCOURT
111	44	281	2	BH609712	HIV18E03	184	44	1.4	803	9	BZ610701	BZ610701	WHAUJ70TR
112	44	294	1	AA493894	nh07b12.s	185	44	1.4	828	6	CA442440	CA442440	UI-H-DP0-
113	44	301	1	AI382825	ta72h10.x	186	44	1.4	830	6	CB307968	CB307968	AGENCOURT
114	44	301	2	BF879334	IL3-ET011	187	44	1.4	830	9	AO781442	AO781442	HS_3117_A
115	44	325	2	BP914587	IL3-UT011	188	44	1.4	834	9	BZ612098	BZ612098	WHAUJ72TR
116	44	337	2	AO959204	CIT-HSP-2	189	44	1.4	847	9	AO781550	AO781550	HS_3176_A
117	44	338	8	N63149	y237e10.s1	190	44	1.4	867	2	BF573285	BF573285	602079678
118	44	352	8	AO092587	HS_3003_B	191	44	1.4	903	3	AO749032	AO749032	HS_5575_A
119	44	353	1	AI367551	gv93e05.x	192	44	1.4	982	1	AL570050	AL570050	AL570050
120	44	365	1	AA654781	nc73g04.s	193	44	1.4	1527	4	CR622010	CR622010	full-len9
121	44	378	9	AO061311	CIT-HSP-2	194	44	1.4	108	2	BE141837	BE141837	IL5-HT011
122	44	378	1	AA653916	nc80a05.s	195	44	1.4	132	2	AA728876	AA728876	nv37c08.r
123	44	381	9	AO239365	CIT-HSP-2	196	44	1.4	132	2	BO012771	BO012771	IL5-GN23
124	44	382	5	BU561388	AGENCOURT	197	44	1.4	151	7	CR543084	CR543084	DKFZP470G
125	44	389	1	AA728990	nv22g06.s	198	44	1.4	154	8	N73776	N73776	z661n08.r
126	44	409	8	H44630	ypl19g10.s1	199	44	1.4	157	1	AI611895	AI611895	ct59f11.x
127	44	412	9	AO605062	HS_2119_B	200	44	1.4	168	1	AA018823	AA018823	z657e09.r
128	44	421	7	CR769029	DKFZP468T	201	44	1.4	196	1	AA478297	AA478297	z445d06.s
129	44	426	8	N20066	yx28g04.s1	202	44	1.4	207	8	F17774	F17774	HSPD02335_H
130	44	435	7	CR546695	DKFZP470L	203	44	1.4	213	3	BG616769	BG616769	60261495
131	44	438	9	BE0059	CIT-HSP-385	204	44	1.4	224	7	CN278030	CN278030	170006001
132	44	447	1	AL713064	DKFZP686C	205	44	1.4	238	1	AM818514	AM818514	RC1-ST027
133	44	448	1	AI469968	c189c03.x	206	44	1.4	246	1	AI886467	AI886467	wn39c10.x
134	44	466	1	AO750507	HS_5521_B	207	44	1.4	252	9	AO286384	AO286384	PC111-79
135	44	466	5	AX117238	AO157238	208	44	1.4	254	1	AA501777	AA501777	ng05d09.s
136	44	466	5	AX117238	AO157238	209	44	1.4	261	2	BA151598	BA151598	ng661n12.
137	44	468	2	BE674703	7e94e04.x	210	44	1.4	265	1	AA478355	AA478355	z445d06.r
138	44	472	8	T84449	ydc32d04.r1	211	44	1.4	277	1	AM848108	AM848108	IL3-CIT021
139	44	473	3	BM983814	UI-CF-DU1	212	44	1.4	282	9	AO378964	AO378964	PC111-16
140	44	483	3	BM314257	1952c03.x	213	44	1.4	294	9	B89891	B89891	CIT-HSP-228
141	44	484	5	AO276534	CITB1-E1-	214	44	1.4	308	1	AI873852	AI873852	wn42e07.x
142	44	485	5	BO432539	AGENCOURT	215	44	1.4	329	7	CN263688	CN263688	170004241
143	44	486	8	N35896	yy28c08.s1	216	44	1.4	339	2	BF184537	BF184537	601842843
144	44	494	1	AI908381	RC-BR170-	217	44	1.4	341	1	AA714110	AA714110	nm05b09.s
145	44	495	8	R66218	yp88a10.r1	218	44	1.4	342	3	BM709383	BM709383	UI-E-CQ1-
146	44	512	2	BE146359	MRO-HT020	219	44	1.4	347	7	CR816152	CR816152	AGENCOURT
147	44	516	2	AI888752	wn36a07.x	220	44	1.4	345	1	AI355886	AI355886	gy51g05.x
148	44	517	2	BE940436	RC3-UT006	221	44	1.4	348	8	R76463	R76463	y463h04.s1
149	44	527	9	AO026637	CIT-HSP-2	222	44	1.4	350	1	AV734149	AV734149	UI-S-CQ1-
150	44	530	7	CN389220	170006003	223	44	1.4	356	9	AO263477	AO263477	AV734149
151	44	538	6	CD691041	EST7564.h	224	44	1.4	368	1	AA834798	AA834798	cd98h01.s
152	44	540	7	CV418654	RC3-UT006	225	44	1.4	368	3	BM667130	BM667130	UI-E-DX0-
153	44	543	9	AO713671	HS_5399_B	226	44	1.4	370	8	T80208	T80208	ydc32e02.r1
154	44	551	1	AA715817	nm25c05.s	227	44	1.4	374	7	CV326131	CV326131	CM4-NN000
155	44	558	6	CB129216	K-EST0178	228	44	1.4	378	1	AV688904	AV688904	AV688904
156	44	581	6	CB218356	n19e04.y	229	44	1.4	391	1	AI633479	AI633479	gq04e01.s
157	44	590	1	AA206418	z651a05.s	230	44	1.4	399	9	AA333017	AA333017	nm05b09.s
158	44	602	9	B70450	CIT-HSP-205	231	44	1.4	400	6	AA633017	AA633017	nm05b09.s
159	44	614	1	AL691931	DKFZP313B	232	44	1.4	400	6	FO1045	FO1045	HSP96B042.S
160	44	614	10	CL246302	HSC_01025	233	44	1.4	407	8	AO202847	AO202847	PC111-43
161	44	627	7	CN267712	170005316	234	44	1.4	412	1	AA962276	AA962276	CO05a02.8
162	44	633	9	AO542374	RPCI-11.3	241	44	1.4	412	1	AI828710	AI828710	tr12d11.x

C 242	43	420	5	BU074474	1m76c02.x	C 315	43	702	9	AQ391927	CITBI-EI-
C 243	43	423	1	AM819125	RC3-ST028	C 316	43	704	3	Bi603876	
C 244	43	426	8	R20234	yg18h12.r1	C 317	43	705	1	AU139904	
C 245	43	428	1	AA829110	o078811.b	C 318	43	705	6	CA447891	UI-H-EIO-
C 246	43	447	1	AA706628	an26d10.b	C 319	43	711	8	CK751218	
C 247	43	450	2	BE145089	CM3-HT019	C 320	43	718	6	CD640244	AGENCOURT
C 248	43	459	1	AL121207	DKE26762C	C 321	43	719	8	BE298219	
C 249	43	459	1	AL121235	DKE26762C	C 322	43	719	8	CK758818	
C 250	43	462	6	CD370252	UI-H-FT1-	C 323	43	727	6	CA440514	UI-H-EDO-
C 251	43	465	1	AM675725	ba53e09.x	C 324	43	733	6	CD640273	AGENCOURT
C 252	43	478	1	AI627168	AI627168	C 325	43	746	2	BE789473	
C 253	43	479	8	H14566	Y125h12.r1	C 326	43	757	2	BG496744	
C 254	43	483	11	CR5959520	Homo sap1	C 327	43	758	5	BX503885	
C 255	43	489	1	AL119649	DKE26761K	C 328	43	762	2	BG776280	60265348
C 256	43	493	2	BP854090	MR2-BN009	C 329	43	763	5	BU170732	
C 257	43	501	7	CR546841	DKE26470B	C 330	43	765	9	BZ600717	
C 258	43	502	9	AQ726636	HS_5411.B	C 331	43	769	9	AQ491214	
C 259	43	504	9	AQ135921	HS_3060.A	C 332	43	771	6	CD110463	
C 260	43	505	9	AQ183779	HS_3199.B	C 333	43	779	6	CD521923	
C 261	43	510	2	BG613206	602640756	C 334	43	815	6	CF597287	
C 262	43	511	5	BX953870	DKE26781A	C 335	43	838	5	BUS02279	
C 263	43	521	5	AQ734876	HS_3051.A	C 336	43	850	8	DR763029	
C 264	43	527	3	BM770275	K-ESt70053	C 337	43	858	2	BE908331	
C 265	43	528	9	AC472287	CITBI-EI-	C 338	43	872	6	CD519236	
C 266	43	529	7	CN388806	170006001	C 339	43	879	6	CB987766	
C 267	43	529	9	AC460087	HS_5123.B	C 340	43	882	2	BE787623	601481733
C 268	43	532	1	AL698441	DKE26686K	C 341	43	892	3	BO230118	
C 269	43	543	2	BG254818	602369246	C 342	43	897	3	BM555408	
C 270	43	544	5	BU076635	Im50d01.y	C 343	43	899	6	CD559035	
C 271	43	547	7	CV571317	o011e07.y	C 344	43	909	5	BO881589	
C 272	43	550	1	AU152518	AI152518	C 345	43	918	5	BX348401	BX348401
C 273	43	550	1	AM003867	WG61d02.x	C 346	43	935	2	BG432843	
C 274	43	550	5	BX501587	DKE26779N	C 347	43	940	6	CD519390	
C 275	43	551	5	BU074758	Im76c02.y	C 348	43	966	5	BO691900	
C 276	43	553	1	AL699840	DKE26686M	C 349	43	970	5	BU161423	
C 277	43	554	1	AW979058	EST391168	C 350	43	973	5	BO718493	
C 278	43	574	2	BE395458	601309981	C 351	43	987	5	BX437681	
C 279	43	576	1	AA601216	no14d11.b	C 352	43	990	3	BM469408	
C 280	43	581	1	AM605171	CV0-PT002	C 353	43	994	1	AL541554	
C 281	43	583	7	CN263464	170004240	C 354	43	1015	3	BM913110	
C 282	43	583	6	AZ519883	RPCI-11-7	C 355	43	1023	3	BM555414	
C 283	43	585	6	CD700137	EST16661	C 356	43	1026	3	BM915472	
C 284	43	590	3	BM751907	K-ESt70028	C 357	43	1029	3	BM471041	
C 285	43	596	3	BO010300	UI-H-EDO-	C 358	43	1031	3	BM915471	
C 286	43	606	5	BX486604	DKE26686K	C 359	43	1044	3	BM915370	AGENCOURT
C 287	43	617	5	BG924293	HNK26-1-C	C 360	43	1052	1	AL540730	
C 288	43	614	5	BX643423	DKE26781K	C 361	43	1058	5	BX436528	
C 289	43	620	2	BX506068	DKE26686D	C 362	43	1060	3	BM913019	
C 290	43	620	2	BE867617	601443133	C 363	43	1063	3	BM807930	
C 291	43	625	5	BX507175	DKE26779C	C 364	43	1093	1	AL534375	
C 292	43	629	6	CD640871	AGENCOURT	C 365	43	1096	5	BX404687	
C 293	43	630	1	AI064786	HA0564.Hu	C 366	43	1099	5	BX404683	
C 294	43	636	10	AG174345	Pan tciog1	C 367	43	1139	3	AL568568	
C 295	43	638	10	AG102516	AG102516	C 368	43	1232	1	BM565405	
C 296	43	639	9	AAQ377619	RPCI11-16	C 369	43	1294	4	BC032375	
C 297	43	640	1	AA012964	2e27c03.r	C 370	43	1701	4	CR623963	
C 298	43	646	10	AG007923	Homo sap1	C 371	43	1706	4	CR619443	
C 299	43	647	3	BQ019639	UI-H-EDO-	C 372	43	1793	4	CR620867	
C 300	43	650	10	AG007924	Homo sap1	C 373	43	1793	4	AF461901	
C 301	43	653	10	AG061520	Pan tciog1	C 374	43	5797	4	CR492333	
C 302	43	656	3	BI334807	602989923	C 375	43	7316	4	HSW805983	
C 303	43	659	3	BI858506	603391746	C 376	43	7316	4	CR857086	
C 304	43	660	3	BM742057	K-ESt70014	C 377	43	121	4	AQ422284	
C 305	43	668	3	BM771614	K-ESt70055	C 378	43	144	2	BG230370	
C 306	43	669	6	CD368504	UI-H-FT1-	C 379	43	145	1	AI492587	
C 307	43	672	9	AQ348932	UI-H-FT1-	C 380	43	151	6	CD642086	
C 308	43	677	10	AG155728	Pan tciog1	C 381	43	183	3	BE788167	
C 309	43	683	10	AG155728	Pan tciog1	C 382	43	230	1	AM196615	
C 310	43	689	3	BM940803	603614259	C 383	43	263	8	W20344	
C 311	43	691	3	BM990097	UI-H-DIO-	C 384	43	263	2	BE092341	
C 312	43	692	9	AQ491216	RPCI-11-2	C 385	43	264	1	AM207358	
C 313	43	692	10	AG141656	Pan tciog1	C 386	43	271	1	AI049508	
C 314	43	701	3	BI333819	602999323	C 387	43	300	9	AZ518434	RPCI-11-4

C 388	42	1.3	316	9	AO042901	CIT-HSP-2	461	42	1.3	617	1	AI133033	AI133033	HAI621	Hu
C 389	42	1.3	341	5	BUS85874	AGENCOURT	462	42	1.3	617	7	CR752362	CR752362	DKFZP459K	
C 390	42	1.3	345	1	AA569631	tm38107.8	463	42	1.3	620	7	CR543489	CR543489	DKFZP459M	
C 391	42	1.3	349	1	BUS65227	AGENCOURT	464	42	1.3	621	10	AG019109	AG019109	Homo sapi	
C 392	42	1.3	365	2	BE007886	BE007886	465	42	1.3	621	10	AG077873	AG077873	Pan trogl	
C 393	42	1.3	369	2	BE007890	BE007890	466	42	1.3	624	9	AO538448	AO538448	RPCI-11-3	
C 394	42	1.3	372	1	AI055937	AI055937	467	42	1.3	651	1	AV762002	AV762002	AV762002	
C 395	42	1.3	373	1	AA702614	AA702614	468	42	1.3	659	10	AG142117	AG142117	Pan trogl	
C 396	42	1.3	378	2	BP914971	BP914971	469	42	1.3	660	10	AG045128	AG045128	Pan trogl	
C 397	42	1.3	378	2	BP914982	BP914982	470	42	1.3	664	5	BX471890	BX471890	DKFZP686A	
C 398	42	1.3	390	1	AM501938	AM501938	471	42	1.3	665	7	CR304726	CR304726	Pan trogl	
C 399	42	1.3	396	8	R02559	Y880C03..81	472	42	1.3	665	10	AG079274	AG079274	Pan trogl	
C 400	42	1.3	404	1	AA211305	AA211305	473	42	1.3	668	5	BX506469	BX506469	DKFZP779E	
C 401	42	1.3	404	3	BM991921	BM991921	474	42	1.3	669	10	AG019043	AG019043	Homo sapi	
C 402	42	1.3	408	1	AI088876	AI088876	475	42	1.3	678	5	BX376229	BX376229	Pan trogl	
C 403	42	1.3	408	1	AA099396	AA099396	476	42	1.3	684	10	AG045683	AG045683	Pan trogl	
C 404	42	1.3	409	1	AI627372	AI627372	477	42	1.3	688	6	CD516136	CD516136	AGENCOURT	
C 405	42	1.3	409	1	AO388413	AO388413	478	42	1.3	689	1	AL704286	AL704286	DKFZP686P	
C 406	42	1.3	411	1	AA806191	AA806191	479	42	1.3	693	10	AG176532	AG176532	Pan trogl	
C 407	42	1.3	419	1	AI524515	AI524515	480	42	1.3	695	5	CR164659	CR164659	full-length	
C 408	42	1.3	430	9	AO089692	AO089692	481	42	1.3	700	6	CA777576	CA777576	IP18C02.Y	
C 409	42	1.3	432	9	AO220470	AO220470	482	42	1.3	705	1	AV760497	AV760497	AV760497	
C 410	42	1.3	433	9	AO017707	CIT-HSP-2	483	42	1.3	707	9	AO262047	AO262047	CITBI-EI-	
C 411	42	1.3	437	9	AO774179	HS_2050-B	484	42	1.3	707	9	AO262047	AO262047	CITBI-EI-	
C 412	42	1.3	438	9	AO061783	CIT-HSP-2	485	42	1.3	718	7	CR745334	CR745334	CR745334	
C 413	42	1.3	441	9	AO383650	RPCI11-14	486	42	1.3	722	7	CR856700	CR856700	DKFZP469H	
C 414	42	1.3	450	6	CD239674	FNPBX02	487	42	1.3	727	10	AG187192	AG187192	Pan trogl	
C 415	42	1.3	455	9	AO359570	HS_5033_A	488	42	1.3	728	1	AV763952	AV763952	AV763952	
C 416	42	1.3	459	7	CA848486	h219d06.Y	489	42	1.3	728	2	BE889588	BE889588	601512509	
C 417	42	1.3	459	5	BX646449	DKFZP781A	490	42	1.3	728	9	AO554491	AO554491	RPCI-11-4	
C 418	42	1.3	460	3	BO314448	QVO-BN004	491	42	1.3	740	6	CR986442	CR986442	AGENCOURT	
C 419	42	1.3	461	3	BO321465	MR3-CT046	492	42	1.3	741	11	CR958354	CR958354	Homo sapi	
C 420	42	1.3	462	5	BX114476	BX114476	493	42	1.3	746	7	CR278648	CR278648	170006003	
C 421	42	1.3	464	1	AI536932	tm87a01.x	494	42	1.3	751	6	CR998596	CR998596	AGENCOURT	
C 422	42	1.3	467	9	AO559196	HS_2069_B	495	42	1.3	751	9	B60698	B60698	CIT-HSP-201	
C 423	42	1.3	468	1	AM090797	xc56d09.x	496	42	1.3	757	1	AI122466	AI122466	AI122466	
C 424	42	1.3	471	9	AO116903	HS_2191_A	497	42	1.3	761	7	CR559036	CR559036	DKFZP468P	
C 425	42	1.3	479	9	B39549	HS-1049-B2-	498	42	1.3	781	10	CG846669	CG846669	HSC 00933	
C 426	42	1.3	480	2	BE140880	UP_527-16	499	42	1.3	784	6	CR962505	CR962505	AGENCOURT	
C 427	42	1.3	486	2	BE144443	MRO-HT016	500	42	1.3	797	2	BG778173	BG778173	60266325	
C 428	42	1.3	491	2	BF740930	QV1-HB003									
C 429	42	1.3	493	5	BX487887	DKFZP686K									
C 430	42	1.3	497	5	BX642370	DKFZP686M									
C 431	42	1.3	499	5	BX482214	DKFZP686J									
C 432	42	1.3	506	9	AO486393	RPCI-11-2									
C 433	42	1.3	507	9	AO191426	HS_3233_A									
C 434	42	1.3	513	10	CU529065	HIV65H12-									
C 435	42	1.3	515	9	AO378058	RPCI-11-1									
C 436	42	1.3	517	2	BE856581	7f64h03.x									
C 437	42	1.3	518	6	AO193220	HS_3062_B									
C 438	42	1.3	519	6	CF142069	UT-HF-CB0									
C 439	42	1.3	524	9	AO345355	RPCI11-13									
C 440	42	1.3	529	9	AO394614	CITBI-EI-									
C 441	42	1.3	535	9	AO395337	CITBI-EI-									
C 442	42	1.3	540	1	AO147274	AO147274									
C 443	42	1.3	542	1	AL712813	DKFZP686E									
C 444	42	1.3	550	7	CR545302	DKFZP459I									
C 445	42	1.3	557	9	AO476584	CITBI-EI-									
C 446	42	1.3	559	5	BU075280	in29e06.Y									
C 447	42	1.3	561	1	AI049598	an38a06.x									
C 448	42	1.3	574	5	BX508959	DKFZP686K									
C 449	42	1.3	579	9	AO052779	RPCI11-49									
C 450	42	1.3	580	3	BP302466	BP302466									
C 451	42	1.3	580	3	AO583761	RPCI-11-4									
C 452	42	1.3	582	3	BP227983	BP227983									
C 453	42	1.3	583	3	AO078254	CIT-HSP-2									
C 454	42	1.3	586	9	AZ517590	RPCI-11-4									
C 455	42	1.3	599	9	AO548881	RPCI-11-3									
C 456	42	1.3	600	5	BX471088	DKFZP686F									
C 457	42	1.3	604	9	AZ757834	HSC 00279									
C 458	42	1.3	609	1	AV763305	AV763305									
C 459	42	1.3	613	9	AZ517246	RPCI-11-7									
C 460	42	1.3	613	9	AO347245	RPCI11-11									

461	42	1.3	617	1	AI133033	AI133033	HAI621	Hu
462	42	1.3	617	7	CR752362	CR752362	DKFZP459K	
463	42	1.3	620	7	CR543489	CR543489	DKFZP459M	
C 464	42	1.3	621	10	AG019109	AG019109	Homo sapi	
C 465	42	1.3	621	10	AG077873	AG077873	Pan trogl	
C 466	42	1.3	624	9	AO538448	AO538448	RPCI-11-3	
467	42	1.3	651	1	AV762002	AV762002	AV762002	
468	42	1.3	659	10	AG142117	AG142117	Pan trogl	
469	42	1.3	660	10	AG045128	AG045128	Pan trogl	
C 470	42	1.3	664	5	BX471890	BX471890	DKFZP686A	
C 471	42	1.3	665	7	CR304726	CR304726	Pan trogl	
C 472	42	1.3	665	10	AG079274	AG079274	Pan trogl	
C 473	42	1.3	668	5	BX506469	BX506469	DKFZP779E	
C 474	42	1.3	669	10	AG019043	AG019043	Homo sapi	
C 475	42	1.3	678	5	BX376229	BX376229	Pan trogl	
C 476	42	1.3	684	10	AG045683	AG045683	Pan trogl	
C 477	42	1.3	688	6	CD516136	CD516136	AGENCOURT	
C 478	42	1.3	689	1	AL704286	AL704286	DKFZP686P	
C 479	42	1.3	693	10	AG176532	AG176532	Pan trogl	
C 480	42	1.3	695	5	CR164659	CR164659	full-length	
481	42	1.3	700	6	CA777576	CA777576	IP18C02.Y	
482	42	1.3	705	1	AV760497	AV760497	AV760497	
C 483	42	1.3	707	9	AO262047	AO262047	CITBI-EI-	
C 484	42	1.3	707	9	AO262047	AO262047	CITBI-EI-	
C 485	42	1.3	718	7	CR745334	CR745334	CR745334	
C 486	42	1.3	722	7	CR856700	CR856700	DKFZP469H	
487	42	1.3	727	10	AG187192	AG187192	Pan trogl	
488	42	1.3	728	1	AV763952	AV763952	AV763952	
C 489	42	1.3	728	2	BE889588	BE889588	601512509	
C 490	42	1.3	728	9	AO554491	AO554491	RPCI-11-4	
C 491	42	1.3	740	6	CR986442	CR986442	AGENCOURT	
C 492	42	1.3	741	11	CR958354	CR958354	Homo sapi	
C 493	42	1.3	746	7	CR278648	CR278648	170006003	
C 494	42	1.3	751	6	CR998596	CR998596	AGENCOURT	
C 495	42	1.3	751	9	B60698	B60698	CIT-HSP-201	
C 496	42	1.3	757	1	AI122466	AI122466	AI122466	
497	42	1.3	761	7	CR559036	CR559036	DKFZP468P	
498	42	1.3	781	10	CG846669	CG846669	HSC 00933	
499	42	1.3	784	6	CR962505	CR962505	AGENCOURT	
C 500	42	1.3	797	2	BG778173	BG778173	60266325	

660 bp	mRNA	linear	EST 07-FEB-2003
BX116028	Scarsa fetal heart JNH19W	Homo sapiens	CDNA clone
BX116028	IMAG5:347458	mRNA sequence.	
IMAG5:347458			

## ALIGNMENTS

RESULT 1  
 Bx116028  
 LOCUS  
 DEFINITION  
 IMGP998C11795 ; IMAGE:347458, mRNA sequence.

ACCESSION  
 Bx116028  
 VERSION  
 Bx116028.1 GI:27880354

KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens (human)

REFERENCE  
 1 (bases 1 to 660)  
 Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,  
 Radelof, U., Schneider, D. and Korn, B.  
 Human Unigeneset - RZP3  
 Unpublished (2003)

TITLE  
 JOURNAL  
 COMMENT  
 Contact: Ina Rolfs  
 RZP Deutsche Ressourcenzentrum fuer Genomforschung GmbH  
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
 RZP; IMGP998C11795.

RZP; IMGP998C11795.  
 RZP; IMGP998C11795.  
 RZP; IMGP998C11795.  
 Human Unigeneset - RZP3 (RZP; IMGP998C11795)  
 http://www.rzpd.de/cloneCards/cgi/972  
 bin/showLib.pl.cgi/response?libNo=972 Contact: Ina Rolfs  
 RZP Deutsche Ressourcenzentrum fuer Genomforschung GmbH  
 Heubenerweg 6, D-14059 Berlin, Germany



```

Db      67 TGCTGCTTTGGGGGACGAGAGGAGCCAGTCTCGGGGGACCCGACGTCTGTGCGCA 126
Qy      163 CAGGGTCCGGGAGTCACTATAGTGGGTTCTAGTCCCATCAAGGCAAAACTCCGGGG 222
Db      127 CAGGGTCCGGGAGTCACTATAGTGGGTTCTAGTCCCATCAAGGCAAAACTCCGGGG 186
Qy      223 AGCCTGGCCCGCTTTTACCTGAGGCTCAAGTTTCCCATCCGGAATAAATGAACGGGTTG 282
Db      187 AGCTGGCCCGCTTTTACCTGAGGCTCAAGTTTCCCATCCGGAATAAATGAACGGGTTG 246
Qy      283 ATCTCCGAGCGGTAACTTCCAGAACTCGATGGGCGAAGGAGGAGGAGGAGTGGCC 342
Db      247 ATCTCCGAGCGGTAACTTCCAGAACTCGATGGGCGAAGGAGGAGGAGGAGTGGCC 306
Qy      343 ACCCAACGTGACTCTCCCGGCTGGAGCCCGGCTTCACTGATCCAGGGGCTGGAGCT 402
Db      307 ACCCAACGTGACTCTCCCGGCTGGAGCCCGGCTTCACTGATCCAGGGGCTGGAGCT 366
Qy      403 CCGGCGGGAGCGAGCGGGGTGGGGGTCTTAGAGAAACCTAACCGGCGCGCTTGGCAG 462
Db      367 CCGGCGGGAGCGAGCGGGGTGGGGGTCTTAGAGAAACCTAACCGGCGCGCTTGGCAG 426
Qy      463 CGCTTAAGCGGAGCGCGGCTCTTGACGCTTGTCCCGGAGTTGGCACTCCAGAG 522
Db      427 CGCTTAAGCGGAGCGCGGCTCTTGACGCTTGTCCCGGAGTTGGCACTCCAGAG 486
Qy      523 GATGGGAGCGGACCTCTGACCTTCCGAGGAGGACCGGTGGAGCCAGGGCGGTGGAG 582
Db      487 GATGGGAGCGGACCTCTGACCTTCCGAGGAGGACCGGTGGAGCCAGGGCGGTGGAG 546
Qy      583 ACAACGATGTGACTCGAGATGGGCTTGGGAGAGATGAGACGAGGAGCGGGGAGCCGCTA 642
Db      547 ACAACGATGTGACTCGAGATGGGCTTGGGAGAGATGAGACGAGGAGCGGGGAGCCGCTA 606
Qy      643 ACGGGGCTCTCTTGTGGCGCCCGCTTCCGAGAGCGCACGTGAGGGTCCCGGGCGGGCT 702
Db      607 ACGGGGCTCTCTTGTGGCGCCCGCTTCCGAGAGCGCACGTGAGGGTCCCGGGCGGGCT 666
Qy      703 CGGT 706
Db      667 CGGT 670

```

RESULT 3  
CPI45448 677 bp mRNA linear EST 06-AUG-2003  
LOCUS UI-HF-C80-asg-a-08-0-UI.r1 NIH\_MGC\_210 Homo sapiens cDNA clone  
ACCESSION IMAGE:30569191 5', mRNA sequence.  
VERSION CPI45448  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 677)  
Bonaldo,M.F., Lennon,G. and Soares,M.B.  
Normalization and subtraction: two approaches to facilitate gene  
discovery  
Genome Res. 6 (9), 791-806 (1996)  
JOURNAL PubMed  
COMMENT Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newcom Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@iowa.edu  
Tissue Procurement: Tim Ratliff  
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa

Clone Distribution: Distribution information can be found at  
http://genome.uiowa.edu/distribution/humanf1.html  
The following repetitive elements were found in this cDNA  
sequence: 129-236, >MIR51NE/MIR  
Seq primer: PYX-5.  
Location/Qualifiers

FEATURES  
source 1..677  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30569191"  
/issue\_type="CNCAP(3)T-225 cell line"  
/lab\_host="DH10B (T1 phage resistant)"  
/clone\_idb="NIH MGC 210"  
/note="Organ: Prostate; Vector: pRT3 Pac; Site 1: Ecor I;  
Site 2: Not I; The library was constructed according  
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
1996. Denatured RNA was size fractionated on a 1% agarose  
gel. First strand cDNA synthesis was primed with oligo-dT  
primer containing a Not I site. Double strand cDNA was  
size selected according to mRNA size fraction, ligated  
with Ecor I adaptor, digested with Not I and then cloned  
directionally into pRT3 Pac vector. The library tag  
sequence located between the Not I site and the polyA tail  
is CCCAC. Tissue was provided by Tim Ratliff."

## ORIGIN

Query Match 19.2%; Score 598; DB 6; Length 677;  
Best Local Similarity 99.8%; Pred. No. 2e-279;  
Matches 648; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 43 GCGCAACGGGAGGAGACACCTGACCCCGGCGGCGCCAGCCCTCGATTGCGAGTAC 102  
Db 7 GCGCAACGGGAGGAGACACCTGACCCCGGCGGCGCCAGCCCTCGATTGCGAGTAC 66  
Qy 103 TGCTGCTTTGGGAGCAGGAGGTGCGAGTCTTGGGAGCACCAGAGTCTGTGCGCA 162  
Db 67 TGCTGCTTTGGGAGCAGGAGGTGCGAGTCTTGGGAGCACCAGAGTCTGTGCGCA 126  
Qy 163 CAGGGTCCGGGAGTCACTATAGTGGGTTCTAGTCCCATCAAGGCAAAACTCCGGGG 222  
Db 127 CAGGGTCCGGGAGTCACTATAGTGGGTTCTAGTCCCATCAAGGCAAAACTCCGGGG 186  
Qy 223 AGCCTGGCCCGCTTTTACCTGAGGCTCAAGTTTCCCATCCGGAATAAATGAACGGGTTG 282  
Db 187 AGCTGGCCCGCTTTTACCTGAGGCTCAAGTTTCCCATCCGGAATAAATGAACGGGTTG 246  
Qy 283 ATCTCCGAGCGGTAACTTCCAGAACTCGATGGGCGAAGGAGGAGGAGTGGCC 342  
Db 247 ATCTCCGAGCGGTAACTTCCAGAACTCGATGGGCGAAGGAGGAGGAGTGGCC 306  
Qy 343 ACCCAACGTGACTCTCCCGGCTGGAGCCCGGCTTCACTGATCCAGGGGCTGGAGCT 402  
Db 307 ACCCAACGTGACTCTCCCGGCTGGAGCCCGGCTTCACTGATCCAGGGGCTGGAGCT 366  
Qy 403 CCGGCGGGAGCGAGCGGGGTGGGGGTCTTAGAGAAACCTAACCGGCGCGCTTGGCAG 462  
Db 367 CCGGCGGGAGCGAGCGGGGTGGGGGTCTTAGAGAAACCTAACCGGCGCGCTTGGCAG 426  
Qy 463 CGCTTAAGCGGAGCGCGGCTCTTGACGCTTGTCCCGGAGTTGGCACTCCAGAG 522  
Db 427 CGCTTAAGCGGAGCGCGGCTCTTGACGCTTGTCCCGGAGTTGGCACTCCAGAG 486  
Qy 523 GATGGGAGCGGACCTCTGACCTTCCGAGGAGGACCGGTGGAGCCAGGGCGGTGGAG 582  
Db 487 GATGGGAGCGGACCTCTGACCTTCCGAGGAGGACCGGTGGAGCCAGGGCGGTGGAG 546  
Qy 583 ACAACGATGTGACTCGAGATGGGCTTGGGAGAGATGAGACGAGGAGCGGGGAGCCGCTA 642  
Db 547 ACAACGATGTGACTCGAGATGGGCTTGGGAGAGATGAGACGAGGAGCGGGGAGCCGCTA 606  
Qy 643 ACGGGGCTCTCTTGTGGCGCCCGCTTCCGAGAGCGCACGTGAGGGTCCCGGGCGGGCT 691

Db 607 ACAGGAGCTCCTCTGCGCGCCGCCGAGAGCGCAGCTCGAGGGTTC 655

RESULT 4  
BM671616/c 685 bp mRNA linear EST 27-FEB-2002  
LOCUS UI-E-CQ1-aggc-1-04-0-UI.s1 UI-E-CQ1 Homo sapiens cDNA clone  
DEFINITION UI-E-CQ1-aggc-1-04-0-UI 3', mRNA sequence.  
ACCESSION BM671616  
KEYWORDS BM671616.1 GI:18981514  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 685)  
AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
PUBMED 8889548  
COMMENT Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@iowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA library preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Gene Distribution: Researchers may obtain clones from Research  
Genetics (www.resgen.com).  
Seq primer: M13 Forward  
POLYA=Yes.

FEATURES  
source Location/Qualifiers  
1..685  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-E-CQ1-aggc-1-04-0-UI"  
/tissue\_type="optic nerve"  
/dev\_stage="adult"  
/lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
/note="Organ: eye; Vector: pRT73-Pac (Pharmacia) with a  
modified polylinker; Site 1: Ecor I; Site 2: Not I;  
UI-E-CQ1 is a normalized cDNA library containing the  
following tissue(s): optic nerve. The library was  
constructed according to Ronaldo, Lennon and Soares,  
Genome Research, 6:791-806, 1996. First strand cDNA  
synthesis was primed with an oligo-dT primer containing a  
Not I site. Double stranded cDNA was ligated to an Ecor I  
adaptor, digested with Not I, and cloned directionally  
into pRT73-Pac vector. The oligonucleotide used to prime  
the synthesis of first-strand cDNA contains a library tag  
sequence that is located between the Not I site and the  
(dT)18 tail. The sequence tag for this library is  
CCATTAAAGT. This library was created for the program, Gene  
Discovery in the Visual System, supported by National Eye  
Institute (NEI).  
TAG\_TISSUE=human optic nerve  
TAG\_LIB=UI-E-CQ1  
TAG\_SEQ=CCATTAAAGT"

ORIGIN  
Query Match 17.0%; Score 531; DB 3; Length 685;  
Best Local Similarity 99.8%; Fred. No. 8.7e-247;  
Matches 651; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
QY 2121 GGCTTCTCTAAGACGCGTAGACGCTTCTCTAGAGAGTTATTCATTGTCCTCCACAG 2180

Db 668 GGCTTCTCTAAGACGCGTAGACGCTTCTCTAGAGAGTTATTCATTGTCCTCCACAG 609

QY 2181 GAGCTAAGAAAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCATT 2240  
Db 608 GAGCTAAGAAAGATTTGAGGTCATGACCTCCCACTGCCGCTCAGGGGCTGACCCATT 549  
QY 2241 AGGAAACCAAGAGGGTGGTGAACCTTACTCTCAGGACTTGGATTCAGTGGCAGACT 2300  
Db 548 AGGAAACCAAGAGGGTGGG-TGAACCTTACTCTCAGGACTTGGATTCAGTGGCAGACT 490  
QY 2301 TGCTGCGGAAAAGGGCTCTCCCAAGCCACCCGAGATGGGGTAAAGAGAGCAGAG 2360  
Db 489 TGCTGCGGAAAAGGGCTCTCCCAAGCCACCCGAGATGGGGTAAAGAGAGCAGAG 430  
QY 2361 GCTTGGGGTGAAGGCCACCTGGTGTAAACAGGACTTCTCTCTGGGGCTTAATT 2420  
Db 429 GCTTGGGGTGAAGGCCACCTGGTGTAAACAGGACTTCTCTCTGGGGCTTAATT 370  
QY 2421 TTGTTCAAGACTAAGCAGAGGTTGAACCTCTTGGAGAGGGCTGGATCCTCTT 2480  
Db 369 TTGTTCAAGACTAAGCAGAGGTTGAACCTCTTGGAGAGGGCTGGATCCTCTT 310  
QY 2481 TAGAGCACTTAATCTTATTTATCCCTGGAATGTGCTGCTGGCCAGTAGAGGGCTGCG 2540  
Db 309 TAGAGCACTTAATCTTATTTATCCCTGGAATGTGCTGCTGGCCAGTAGAGGGCTGCG 250  
QY 2541 TTGAGCACTCTCCCTGACCCCGCGCTGCCCGCCCTCCGGGGTAAATGTGCACTTATGCG 2600  
Db 249 TTGAGCACTCTCCCTGACCCCGCGCTGCCCGCCCTCCGGGGTAAATGTGCACTTATGCG 190  
QY 2601 CCACAGAGGTTTGAAGCAATGAGCTGAGACTGGGTGAATGAATGAACAGCTTAATT 2660  
Db 189 CCACAGAGGTTTGAAGCAATGAGCTGAGACTGGGTGAATGAATGAACAGCTTAACTT 130  
QY 2661 GGGATTTAAGAGCTTTTAAAGATATATCTCTGAAAGAAAATAGATGATGACCAAG 2720  
Db 129 GGGATTTAAGAGCTTTTAAAGATATATCTCTGAAAGAAAATAGATGATGACCAAG 70  
QY 2721 CGTGTACTAATGAAGCTGTTATTTTAAATGAAGAGCTGGGCGATGAACTCA 2772  
Db 69 CGTGTACTAATGAAGCTGTTATTTTAAATGAAGAGCTGGGCGATGAACTCA 18

RESULT 5  
LOCUS BM707056 578 bp mRNA linear EST 28-FEB-2002  
DEFINITION UI-E-CR1-adx-b-12-0-UI.r1 UI-E-CR1 Homo sapiens cDNA clone  
VERSION BM707056  
KEYWORDS UI-E-CR1-adx-b-12-0-UI 5', mRNA sequence.  
SOURCE BM707056.1 GI:19020314  
ORGANISM Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 578)  
AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.  
TITLE Normalization and subtraction: two approaches to facilitate gene  
discovery  
JOURNAL Genome Res. 6 (9), 791-806 (1996)  
PUBMED 8889548  
COMMENT Contact: Soares, MB  
Coordinated Laboratory for Computational Genomics  
University of Iowa  
375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
Tel: 319 335 8250  
Fax: 319 335 9565  
Email: bento-soares@iowa.edu  
Tissue Procurement: Dr. Gregg Hageman  
cDNA library preparation: Dr. M. Bento Soares, University of Iowa  
CDNA library Arrayed by: Dr. M. Bento Soares, University of Iowa







OY	2142	TAGCCCTCTCAGCAGAGTTTATCCATTCTGTCCTCCAGACGCTAGAAAGA	TTTGAAG	2201
Db	121	CAGCCCTCTCAGCAGAGTTTATCCATTCTGTCCTCCAGACGCTAGAAAGA	TTTGAAG	180
OY	2202	TCATGACCTCCACCTGCGCTCAGGGGCTGACCCCTATTAGAAA	CCAAAGGGCTGGT	2261
Db	181	TCATGACCTCCACCTGCGCTCAGGGGCTGACCCCTATTAGAAA	CCAAAGGGCTGGT	240
OY	2262	TGAACCTACTCTCAGGACTTGGATCCAGTGGCA	CACTTGCTGCGAAAAAGGCTCTC	2321
Db	241	TGAACCTACTCTCAGGACTTGGATCCAGTGGCA	CACTTGCTGCGAAAAAGGCTCTC	300
OY	2332	CCGACGCACCGGAGATGGGGGTAAAGAGAAAGACAGAGGCTTGGGGTA	AGGGCCACCTGG	2381
Db	301	CCGACGCACCGGAGATGGGGGTAAAGAGAGAGAGGCTTGGGGTA	AGGGCCACCTGG	360
OY	2382	TGTTTAAACAGGCACTTCTCTCTCTCGGGCTTATTTTGTTCAGAACT	AGACAG	2441
Db	361	TGTTTAAACAGGCACTTCTCTCTCTCTCGGGCTTATTTTGTTCAGAACT	AGACAG	420
OY	2442	TGTTTGAACCTCTCTTTCAGAGAGGCTGGGAATCTCTTTAGAGCACT	TAAATCTTATTTA	2501
Db	421	TGTTTGAACCTCTCTTTCAGAGAGGCTGGGAATCTCTTTAGAGCACT	TAAATCTTATTTA	480
OY	2502	TCCCTCGAATCTGGTGCT	2521	
Db	481	TCCCTCGAATCTGGTGCT	500	

```

/dev stage="fetal and adult"
/lab_host="PH10B (Life Technologies) (T1 phage resistant)"
/clone_1lb="UI-E-Ex1"
/note="Organ: eye; Vector: pT7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: EcoR I; Site_2: Not I;
UI-E-Ex1 is a subcloned cDNA library constructed
according to Bonaldo, Lennon and Soares, Genome Research,
6:791-806, 1996. First strand cDNA synthesis was primed
with an oligo-dT primer containing a Not I site. Double
stranded cDNA was ligated to an EcoR I adaptor, digested
with Not I, and cloned directionally into pT7T3-Pac
vector. The oligonucleotide used to prime the synthesis of
first-strand cDNA contains a library tag sequence that is
located between the Not I site and the (dr)18 tail. The
sequence tags for this library are: fetal eyes,
AGAGTCAGACA; lens, CGATTAGCGA; eye anterior segment,
AATGGCCGCA; optic nerve, CCATTAAATG; retina, CCGCG; Retina
Poreal and Macular, GTCC; RPB and Choroid, ACCCA. This
library was created for the program, Gene Discovery in the
Visual System, supported by National Eye Institute (NEI)."

```

Query Match	14.4%	Score	450;	DB 3;	Length	655;	
Best Local Similarity	99.4%	Pred. No.	2.6e-207;				
Matches 650; Conservative	0;	Mismatches	4;	Indels	0;	Gaps	0;

OY	526	GGGAGCCGCAACCCCTACACTTCGACAGGGAGGCAACGCTGAGAGCCAGAGGCGGTGACAGACACA	585
Db	2	GGGGACCGCACCTCTCACACTTCGACAGGGAGGCAACGCTGAGAGCCAGAGGCGGTGACAGACACA	61
OY	586	CGACGTGTGACTCGAGATGCGCCTTGAGGAGAGATGGAAGAGGAGCGGGAGACCGCTAAACG	645
Db	62	CGACGTGTGACTCGAGATGCGCCTTGAGGAGAGATGGAAGAGGAGCGGGAGACCGCTAAACG	121
OY	646	GGGCTCCTCTCTGCGCGCGCCCGTCCGACAGGCGCACTCGAGGGTCCCGGGCGGGCTTCG	705
Db	122	GGGCTCCTCTCTGCGCGCGCCCGTCCGACAGGCGCACTCGAGGGTCCCGGGCGGGCTTCG	181
OY	706	TGAGCGTTTGGCGGGTAGCGCGGAGGATGCACGAGCACTGAAGAAGCTTTCGTGCGCGCGG	765
Db	182	TGAGCGTTTGGCGGGTAGCGCGGAGGATGCACGAGCACTGAAGAAGCTTTCGTGCGCGCGG	241
OY	766	CCCAAGGCGCGGAGATGGGGGTTTAACTCAATCTCTGCGCGCTTGAAGGGAGGCTTAAACGGGCG	825
Db	242	CCCAAGGCGCGGAGATGGGGGTTTAACTCAATCTCTGCGCGCTTGAAGGGAGGCTTAAACGGGCG	301
OY	826	CGGGCGGCGCGGCGCGGAGCGCGGAGCCCAACCGCGATGCGGAGGAGAGTGAAGAAGCGGTTCG	885
Db	302	CGGGCGGCGCGGCGCGGAGCGCGGAGCCCAACCGCGATGCGGAGGAGAGTGAAGAAGCGGTTCG	361
OY	886	TGGAACGGGCTCAACAAGACGACTGCGGTCTAACACACACTGATGATCTGACCGTCCGATGCT	945
Db	362	TGGAACGGGCTCAACAAGACGACTGCGGTCTAACACACACTGATGATCTGACCGTCCGATGCT	421
OY	946	CGGCGGACTTCGCAGAACTTCGCGGCGAGAGCTTGCAGAAAGACGCGCCAGAAAGCGCCAGAGAC	1005
Db	422	CGGCGGACTTCGCAGAACTTCGCGGCGAGAGCTTGCAGAAAGAGCGCGCCAGAAAGCGCCAGAGAC	481
OY	1006	TGCGCGGTCTCAACTGCGGCGCGGCTCTACTGCTGCTGCGGACCGGGGCGCTTGGCGGCGG	1065
Db	482	TGCGCGGTCTCAACTGCGGCGCGGCTCTACTGCTGCTGCGGACCGGGGCGCTTGGCGGCGG	541
OY	1066	ACGAGCGCGCGGAGTTGAGCGGCTCTGAGGTGAGCTTCTCGGAGCTGCTGAACCTGCTGG	1125
Db	542	ACGAGCGCGCGGAGTTGAGCGGCTCTGAGGTGAGCTTCTCGGAGCTGCTGAACCTGCTGG	601
OY	1126	AAAGCGACATGCGACGCTCGCTGAGGTGGGCGCCGCGTTCCGCTGACGCGC	1179
Db	602	AAAGCGACATGCGACGCTCGCTGAGGTGGGCGCCGCGTTCCGCTGACGCGC	655

FEATURES	source	location/Qualifiers
		1..655
		/organism="Homo sapiens"
		/mol_type="mRNA"
		/db_xref="caxon:9606"
		/clone="UI-B-E1-a]-z-e-14-0-UI"
		/tissue_type="fetal eyes, lens, eye anterior segment,"
		optic nerve, retina, Retina Foveal and Macular, RPE and
		Choroid"

RESULT 8  
AW302149/C

LOCUS AM302149 476 bp mRNA linear EST 18-JUN-2000  
DEFINITION x801f06.x1 NCI\_CGAP Kid11 Homo sapiens cDNA clone IMAGE:2768387 3'  
similar to contatins TAK1 repetitive element ;, mRNA sequence.  
ACCESSION AM302149  
VERSION AM302149.1 GI:6711826  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
1 (bases 1 to 476)  
AUTHORS NCI-CGAP  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA library Preparation: M. Bento Soares, Ph.D.  
cDNA library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
www.bio.llnl.gov/dbrr/image/image.html

FEATURES  
source  
Possible reversed clone: polyT not found  
Seq primer: -40UP from Gibco  
High quality sequence stop: 457.  
Location/Qualifiers  
1..476  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2768387"  
/lab\_host="DH10B"  
/clone\_1ib="NCI CGAP Kid11"  
/note="Organ: kidney; Vector: pT773D-Pac (Pharmacia) with  
a modified polylinker; Site 1: Not 1; Site 2: Eco RI;  
Plasmid DNA from the normalized library NCI CGAP Kid3 was  
prepared, and as circles were made in vitro. Following HAP  
purification, this DNA was used as tracer in a subtractive  
hybridization reaction. The driver was PCR-amplified cDNAs  
from a pool of 5,000 clones made from the same library  
(cloneids 1322376-1323911, 1456007-1456775, and  
1500552-1502855). Subtraction by Bento Soares and M.  
Facina Bonaldo."

ORIGIN  
Query Match 13.7%; Score 427; DB 1; Length 476;  
Best Local Similarity 100.0%; Pred. No. 4.2e-196;  
Matches 427; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 0;  
716 CCGTAGCGCCGAGCAGTACCGAGCCATTAAGAGCTTGTGCGCGCGCCCAAGCGG 775  
Db 476 CCGTAGCGCCGAGCAGTACCGAGCCATTAAGAGCGTTGTGCGCGCGCCCAAGCGG 417  
Qy 776 GGAATGGGGGTTAGCCATCTCTGCGCGCTTGAAGGGGAGGCTTAACTGGGCGCGGCGCG 835  
Db 416 GGAATGGGGGTTAGCCATCTCTGCGCGCTTGAAGGGGAGGCTTAACTGGGCGCGGCGG 357  
Qy 836 GGGCCCGCCGAGCCCACTGCGAGTGGCGAGGAGGAGTCAAGAGCGCTCTGAACGGGCT 895  
Db 356 GGGCCCGCCGAGCCCACTGCGAGTGGCGAGGAGGAGTCAAGAGCGCTCTGAACGGGCT 297  
Qy 896 CAACAAGAGAGCTGCTGTACCACTGTGTGTGACCGGTGCGGTGCTGGCGGAGCTC 955  
Db 296 CAACAAGAGAGCTGCTGTACCACTGTGTGTGACCGGTGCGGTGCTGGCGGAGCTC 237  
Qy 956 GCAGAACTGCGCGAGAGCTGCAAAAGACGCCGAGAGGCGCAGAGAGCTGGCGGTGTC 1015  
Db 236 GCAGAACTGCGCGAGAGCTGCAAAAGACGCCGAGAGGCGCAGAGAGCTGGCGGTGTC 177

Qy 1016 CACCTGCCCCCGGCTGACTGTGTGTGCTGCGGAGACCGGGGCTTGGCCCGCGAGAGCGGCG 1075  
Db 176 CACCTGCCCCCGGCTGACTGTGTGTGCTGCGGAGACCGGGGCTTGGCCCGCGAGAGCGGCG 117  
Qy 1076 CGAGTTGAGGCGGCTCTGGTGTGCTTCTCGGAGCTGTGACCTGTGGAAGCGGAGCAT 1135  
Db 116 CGAGTTGAGGCGGCTCTGGTGTGCTTCTCGGAGCTGTGACCTGTGGAAGCGGAGCAT 57  
Qy 1136 GCGAGCGC 1142  
Db 56 GCGAGCGC 50

RESULT 9  
AI375213/c 455 bp mRNA linear EST 18-MAR-1999  
LOCUS tc10f06.x1 Soares NHMPL\_51 Homo sapiens cDNA clone IMAGE:2063459  
DEFINITION 3', mRNA sequence.  
ACCESSION AI375213  
VERSION AI375213.1 GI:4175203  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 455)  
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
JOURNAL Unpublished (1997)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-remail.nih.gov  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 587 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 454.  
Location/Qualifiers  
1..455  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2063459"  
/clone\_type="Pooled human melanocyte, fetal heart, and  
pregnant uterus"  
/lab\_host="DH10B"  
/clone\_1ib="Soares NHMPL\_51"  
/note="Organ: mixed (see below); Vector: pT773D-Pac  
(Pharmacia) with a modified polylinker; Site 1: Not 1;  
Site 2: Eco RI; Equal amounts of plasmid DNA from three  
normalized libraries (melanocyte 2NBH, pregnant uterus  
NBHPL, and fetal heart NBH19W) were mixed, and as circles  
were made in vitro. Following HAP purification, this DNA  
was used as tracer in a subtractive hybridization  
reaction. The driver was PCR-amplified cDNAs from pools of  
5,000 clones made from the same 3 libraries. The pools  
consisted of I.M.A.G.E. clones 260232-265223,  
340488-345479, and 484488-489479."

ORIGIN  
Query Match 13.2%; Score 413; DB 1; Length 455;  
Best Local Similarity 100.0%; Pred. No. 2.8e-189;  
Matches 413; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 0;  
2362 CTTGGGGTGAAGGCACTGTGTGTTAAACAGGCACTTCTCTTCTGCGGCTTAATTT 2421  
Db 415 CTTGGGGTGAAGGCACTGTGTGTTAAACAGGCACTTCTCTTCTGCGGCTTAATTT 356  
Qy 2422 TGTTCAGAACTTGAACCAAGTGTGTTGAACCTCTCTTTCAGAGAGGCTGGAAATCTCTTT 2481  
Db 355 TGTTCAGAACTTGAACCAAGTGTGTTGAACCTCTCTTTCAGAGAGGCTGGAAATCTCTTT 296

QY 2482 AGAGCACTTAATCTATTATCCCTGGAATGCGTGGCCAGTAGAGGGCTGCT 2541  
 DB 235 AGAGCACTTAATCTATTATCCCTGGAATGCGTGGCCAGTAGAGGGCTGCT 236  
 QY 2542 TTGGACGCTCCCTGACCCCGCGCTGCGCCCTCCGCGGTAATGTGGCACTTACTGACC 2601  
 DB 235 TTGGACGCTCCCTGACCCCGCGCTGCGCCCTCCGCGGTAATGTGGCACTTACTGACC 176  
 QY 2602 CACAGAGGTTTGGACCAATCAGCTCTGAGACTGGGTTGAATGTAAACGCTTTAACTTG 2661  
 DB 175 CACAGAGGTTTGGACCAATCAGCTCTGAGACTGGGTTGAATGTAAACGCTTTAACTTG 116  
 QY 2662 GATTATTAAGAGCTTTTAAAGGTAAATATCCCTGGAAGAAATGACGTAAACCAACG 2721  
 DB 115 GATTATTAAGAGCTTTTAAAGGTAAATATCCCTGGAAGAAATGACGTAAACCAACG 56  
 QY 2722 GTGACTATGAAGCTGTATTATTTAATAAAGACGCTGGGCCATGAACCTCAT 2774  
 DB 55 GTGACTATGAAGCTGTATTATTTAATAAAGACGCTGGGCCATGAACCTCAT 3

RESULT 10  
 BM673230/c 480 bp mRNA linear EST 27-FEB-2002  
 LOCUS UI-E-CRI-idx-b-12-0-UI-81 UI-E-CRI Homo sapiens cDNA clone  
 DEFINITION UI-E-CRI-idx-b-12-0-UI 3', mRNA sequence.  
 ACCESSION BM673230  
 VERSION BM673230.1 GI:16983128  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 AUTHORS 1 (bases 1 to 480)  
 TITLE Ronaldo, M. F., Lennon, G. and Soares, M. B.  
 COMMENT Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 PUBMED 8889548  
 CONTACT: Soares, MB  
 COORDINATED LABORATORY for Computational Genomics  
 UNIVERSITY of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Tissue Procurement: Dr. Gregg Hageman  
 CDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Researchers may obtain clones from Research  
 Genetics (www.resgen.com).  
 Seq primer: M13 Forward  
 POLYA=yes

# FEATURES

Source location/Qualifiers  
 1..480  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-E-CRI-idx-b-12-0-UI"  
 /issue\_type="eye anterior segment"  
 /dev\_stage="adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /note="Organ: eye; Vector: pT73-Pac (Pharmacia) with a  
 modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 UI-E-CRI is a normalized cDNA library containing the  
 following tissue(s): eye anterior segment. The library was  
 constructed according to Ronaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand cDNA  
 synthesis was primed with an oligo-dT primer containing a

## ORIGIN

Not I site. Double stranded cDNA was ligated to an EcoR I  
 adaptor, digested with Not I, and cloned directionally  
 into pT73-Pac vector. The oligonucleotide used to prime  
 the synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the  
 (dT)18 tail. The sequence tag for this library is  
 AATGCCGAT. This library was created for the program, Gene  
 Discovery in the Visual System, supported by National Eye  
 Institute (NEI).  
 TAG: TISSUE-human eye anterior segment  
 TAG: LIB=UI-E-CRI  
 TAG\_SEQ=AATGCCGAT"

Query Match 13 2%; Score 412; DB 3; Length 480;  
 Best Local Similarity 99.8%; Pred. No. 8.7e-189;  
 Matches 462; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2310 AAAAGGCTCTCCCGACGACCCCGAGATGGGGTAAAGAGACAGAGGCTTGGGCT 2369  
 DB 480 AAAAGGCTCTCCCGACGACCCCGAGATGGGGTAAAGAGACAGAGGCTTGGGCT 421  
 QY 2370 AGGGCCACTGTGTATTAACAGGCACTTCTCTCTCTGCGGCTTATTTTGTTCAGA 2429  
 DB 420 AGGGCCACTGTGTATTAACAGGCACTTCTCTCTCTGCGGCTTATTTTGTTCAGA 361  
 QY 2430 ACTAAGCAGAGTGTATTAACCTCTTTGACAGAGGCTGGGAATCCTTTAGAGCACT 2489  
 DB 360 ACTAAGCAGAGTGTATTAACCTCTTTGACAGAGGCTGGGAATCCTTTAGAGCACT 301  
 QY 2490 TAATCTATTATCCCTGGAATGCGTGTGCGCAAGTAGAGGCTGCTTGGCAGC 2549  
 DB 300 TAATCTATTATCCCTGGAATGCGTGTGCGCAAGTAGAGGCTGCTTGGCAGC 241  
 QY 2550 TCCCTGACCCCGCGCTGCGCCCTCCGCGGTAATGTGCATTACTGCCCAAGAG 2609  
 DB 240 TCCCTGACCCCGCGCTGCGCCCTCCGCGGTAATGTGCATTACTGCCCAAGAG 181  
 QY 2610 TTTTAGAGCAATCAGCTCTGAGACTGGGTTGAATGTAAACGCTTTAATCTGGGANTTAA 2669  
 DB 180 TTTTAGAGCAATCAGCTCTGAGACTGGGTTGAATGTAAACGCTTTAATCTGGGANTTAA 121  
 QY 2670 GAAGCTTTTAAAGGTAAATATCTCTGAAAGAAATGACGTAAACAGAGCTGTACTA 2729  
 DB 120 GAAGCTTTTAAAGGTAAATATCTCTGAAAGAAATGACGTAAACAGAGCTGTACTA 61  
 QY 2730 TGAAGCTGTATTTTAAATAAAGACGCTGGCCATGAATCTCA 2772  
 DB 60 TGAAGCTGTATTTTAAATAAAGACGCTGGCCATGAATCTCA 18

RESULT 11  
 BQ682843 972 bp mRNA linear EST 15-JUL-2002  
 LOCUS BQ682843  
 DEFINITION AGENCOURT 8495014 NIH\_MGC\_112 Homo sapiens cDNA clone IMAGE:6295149  
 5', mRNA sequence.  
 ACCESSION BQ682843  
 VERSION BQ682843.1 GI:21795522  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 AUTHORS 1 (bases 1 to 972)  
 TITLE NIH-MGC http://mgi.nci.nih.gov/.  
 JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)  
 COMMENT Unpublished (1999)  
 CONTACT: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: DCTD/DTF  
 CDNA Library Preparation: Rubin Laboratory  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/BLNL at:  
http://image.llnl.gov  
Plate: L1CM2500 row: 0 column: 22  
High quality sequence stop: 463.  
Location/Qualifiers

## FEATURES

source

1..972  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6295149"  
/tissue\_type="melanotic melanoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_112"  
/note="Organ: skin; Vector: pORF7; Site: 1: XhoI; Site: 2:  
EcoRI; cDNA made by oligo-dr priming. Directionally cloned  
into EcoRI/XhoI sites using the following 5' adaptor:  
GGCAGGAG(G). Library constructed by Ling Hong in the  
laboratory of Gerald M. Rubin (University of California,  
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and  
Superscript II RT (Life Technologies). Note: this is a  
NIH\_MGC library."

## ORIGIN

Query Match 12.7%; Score 395; DB 5; Length 972;  
Best Local Similarity 100.0%; Pred. No. 1.6e-180; Indels 0; Gaps 0;  
Matches 395; Conservative 0; Mismatches 0;

1899 AGTTCGATGCCACCCCTGCGACGCTGCGCCCTTTCGCGTGGACAGTTGAAAAGGT 1958  
1 AGTTGGTGTCCACCCCTGCGACGCTGCGCCCTTTCGCGTGGACAGTTGAAAAGGT 60  
1959 GGGTGGGATGAGTGAAGTTTGAAGAGGACCGCTGTTGCTTATGTCGTTCTGTT 2018  
61 GGGTGGGATGAGTGAAGTTTGAAGAGGACCGCTGTTGCTTATGTCGTTCTGTT 120  
2019 TCCCGGACAGAAAATTCGATCAATGTCAGAGCTTTTATTCCTTAATCTTTCAGG 2078  
121 TCCCGGACAGAAAATTCGATCAATGTCAGAGCTTTTATTCCTTAATCTTTCAGG 180  
2079 GCTTAATTTTGAAGAGTGTCTCTGAGAGCACTTATCAAAAGGCTTCTCTAAGACGG 2138  
181 GCTTAATTTTGAAGAGTGTCTCTGAGAGCACTTATCAAAAGGCTTCTCTAAGACGG 240  
2139 CTACAGCCCTTCTTACAGAGTTTATCCATTGTCTCCCAAGAGCTTGAAGAAGATTGG 2198  
241 CTACAGCCCTTCTTACAGAGTTTATCCATTGTCTCCCAAGAGCTTGAAGAAGATTGG 300  
2199 AGGTGATGACCTCCGACGCGCTGACAGGGGCTGACCTATTGAAGAAACCAAGAGGGTTG 2258  
301 AGGTGATGACCTCCGACGCGCTGACAGGGGCTGACCTATTGAAGAAACCAAGAGGGTTG 360  
2259 GGTGAACCTACTCTACAGCACTTGATCCAGTGC 2293  
361 GGTGAACCTACTCTACAGCACTTGATCCAGTGC 395

RESULT 12  
BM562937 1376 bp mRNA linear EST 20-FEB-2002  
LOCUS BM562937  
DEFINITION AGENCOURT\_6566935 NIH\_MGC\_88 Homo sapiens cDNA clone IMAGE:5736787  
5', mRNA sequence.  
ACCESSION BM562937  
VERSION BM562937.1 GI:18809436  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 1376)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/BLNL at:  
http://image.llnl.gov  
Plate: L1LM12745 row: n column: 20  
High quality sequence stop: 425.  
Location/Qualifiers

## FEATURES

source

1..1376  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5736787"  
/tissue\_type="duodenal adenocarcinoma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH\_MGC\_88"  
/note="Organ: small intestine; Vector: pCMV-SPORT6;  
Site 1: NotI; Site 2: SalI; Cloned unidirectionally;  
oligo-dr primed. Average insert size 1.767 kb. Library  
enriched for full-length clones and constructed by Life  
Technologies. Note: this is a NIH\_MGC library."

## ORIGIN

Query Match 11.1%; Score 348; DB 3; Length 1376;  
Best Local Similarity 100.0%; Pred. No. 1.3e-157; Indels 0; Gaps 0;  
Matches 348; Conservative 0; Mismatches 0;

453 CCTTGGCAGCGCTTAAGGCGGAGGCGCGGCTCTGCAAGCTGCTTGGCCCGAGATTGGC 512  
1 CCTTGGCAGCGCTTAAGGCGGAGGCGCGGCTCTGCAAGCTGCTTGGCCCGAGATTGGC 60  
513 ACCCAGGAGGATGGGGAGCGCACCCCTCAGCTTGGCAGGGAGCCACGTTGAGGCGAAGG 572  
61 ACCCAGGAGGATGGGGAGCGCACCCCTCAGCTTGGCAGGGAGCCACGTTGAGGCGAAGG 120  
573 CGGTGACAGACACGACGCTGTGACTCGAGTGCCTGCGGAGGATGACGAGGAGCGG 632  
121 CGGTGACAGACACGACGCTGTGACTCGAGTGCCTGCGGAGGATGACGAGGAGCGG 180  
633 GGGACCGCTAACGGGGCTCCTCTGCGCGCCCGTCCGCAAGGCGCACCTGAGAGGTCC 692  
181 GGGACCGCTAACGGGGCTCCTCTGCGCGCCCGTCCGCAAGGCGCACCTGAGAGGTCC 240  
693 CGGGGCGGCTCGTGGAGAGTTGGCGGATGCGCGAGGAGCTACAGGACCTAGAGAGGCT 752  
241 CGGGGCGGCTCGTGGAGAGTTGGCGGATGCGCGAGGAGCTACAGGACCTAGAGAGGCT 300  
753 TGTGTCGCGCGCGCCCAAGGCGGGATGGGGATTAGCCACTCTTGGC 800  
301 TGTGTCGCGCGCGCCCAAGGCGGGATGGGGATTAGCCACTCTTGGC 848

RESULT 13  
W81663 595 bp mRNA linear EST 17-OCT-1996  
LOCUS W81663  
DEFINITION zdb5906.r1 Soares fetal heart NBH1919W Homo sapiens cDNA clone  
IMAGE:347458 5', mRNA sequence.  
ACCESSION W81663  
VERSION W81663.1 GI:1392291  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
REFERENCE 1 (bases 1 to 595)  
AUTHORS Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,



```

Db      120 GNGCCCTGGGAGGATGAGGAGGAGGAGGAGCCGCTAACGGGGCTCCCTTCGCGG 61
QY      662 CCCCCTCCGACAGGCGCACGTCGAGGCTCCCGGGCGGGCTCCGTCGACG 711
Db      60 CCCCCTCCGACAGGCGCACGTCGAGGCTCCCGGGCGGGCTCCGTCGACG 11

RESULT 15
LOCUS   BX645932              728 bp    mRNA    linear    EST 04-SEP-2003
DEFINITION DKEF2p781B1952.r1.781 (synonym: h1cc4) Homo sapiens cDNA clone
ACCESSION BX645932
VERSION   BX645932
KEYWORDS  BX645932.1 GI:34480265
SOURCE    EST.
ORGANISM  Homo sapiens (human)
MIMIS     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 728)
AUTHORS   Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Mewes,H.W., Weil,B., Amd,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.
EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I.,
Wellenreuther,R., et al.)
JOURNAL   Unpublished (2003)
COMMENT   Contact: MIPS
MIPS      Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
This is from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
No 31 sequence available.
Please contact the RZPD: Resequenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
1..728
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp781B1952"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="781 (synonym: h1cc4)"
/note="Vector: pSport1_Sfi; Site_1: SfiIA; Site_2: SfiIB;
cDNA-collection"

ORIGIN
Query Match      5.8%; Score 180; DB 5; Length 728;
Best Local Similarity 100.0%; Pred. No. 9.8e-76;
Matches 180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2693 CTCCTGAAAAGAAAAGAGCTAACCAAGGCTACTATGAAAGCTTTATTATTAAG 2752
Db      21 CTCCTGAAAAGAAAAGAGCTAACCAAGGCTACTATGAAAGCTTTATTATTAAG 80
QY      2753 AAGCGTGGGCGCATGATCATCTGCGCAATGAGTCAACATGATCTTTATGAGATA 2812
Db      81 AAGCGTGGGCGCATGATCATCTGCGCAATGAGTCAACATGATCTTTATGAGATA 140
QY      2813 CTTAAGTTACTAATATATATATTTCTATCTTCTGAAAGTTGATGCTTCCCGCCCCC 2872
Db      141 CTTAAGTTACTAATATATATATTTCTATCTTCTGAAAGTTGATGCTTCCCGCCCCC 200

RESULT 16
LOCUS   W81222              449 bp    mRNA    linear    EST 17-OCT-1996

```

```

DEFINITION zdbse06.b1 Soares fetal heart NBH119W Homo sapiens cDNA clone
IMAGE:347458 3', mRNA sequence.
W81222
VERSION   W81222.1 GI:1392242
KEYWORDS  EST.
SOURCE    Homo sapiens (human)
ORGANISM  Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
REFERENCE 1 (bases 1 to 449)
AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,
Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marx,M.,
Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F.,
Trevaaskis,B., Waterson,R., Williamson,A., Woldmann,P. and
Wilson,R.
JOURNAL   The Mashu-Merck EST Project
Unpublished (1995)
COMMENT   Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 584 Std Error: 0.00
Seq primer: mob.REGA+ET
High quality sequence stop: 76.
FEATURES
source
1..449
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1272833"
/db_xref="taxon:9606"
/clone="IMAGE:347458"
/sex="unknown"
/dev_stage="19 weeks"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares_fetal_heart_NBH119W"
/note="Organ: heart; Vector: pTV73D (Pharmacia) with a
modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCATCTTTTATTTTATTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pTV73 vector
(Pharmacia). Library went through one round of
normalization to a Cot = 5. Library constructed by
M.Patino Bernaldo. This library was constructed from the
same fetus as the fetal lung library, Soares fetal lung
NBH119W."

ORIGIN
Query Match      4.6%; Score 144; DB 8; Length 449;
Best Local Similarity 100.0%; Pred. No. 3.5e-58;
Matches 144; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2631 GACTGGGTTGAGATGTAACGCTTAACTTGGGATTTAAGAACTTTAAAGTAATA 2690
Db      145 GACTGGGTTGAGATGTAACGCTTAACTTGGGATTTAAGAACTTTAAAGTAATA 86
QY      2691 TCCTGAAAAGAAAAGAGCTAACCAAGGCTACTATGAAAGCTTTATTTATTAATA 2750
Db      85 TCCTGAAAAGAAAAGAGCTAACCAAGGCTACTATGAAAGCTTTATTTATTAATA 26
QY      2751 AGAAGCTGGGCGCATGAACTGATA 2774
Db      25 AGAAGCTGGGCGCATGAACTGATA 2

RESULT 17
LOCUS   AG013685

```



```

LOCUS      AG013685              709 bp    DNA          linear    GSS 16-FEB-2005
DEFINITION Homo sapiens genomic DNA, 21q region, clone: B137B7N13, genomic
survey sequence.
ACCESSION  AG013685
VERSION    AG013685.2  GI:55788901
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE  1
AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE     Homo sapiens genomic DNA, chromosome 21q
JOURNAL   Published Only in Database (1998)
REFERENCE  2 (bases 1 to 709)
AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE     Direct Submission
JOURNAL   Submitted (09-SEP-1998) Masahira Hattori, RIKEN Genomic Sciences
Center, RIKEN Yokohama Institute, Yokohama Research Promotion
Division, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa,
230-0045, Japan (E-mail:hattori@gscc.riken.jp, Tel:81-45-503-9111,
Fax:81-45-503-9113)
COMMENT   On Nov 16, 2004 this sequence version replaced gi:3551633.
FEATURES   Location/Qualifiers
            source          1..709
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="21"
                        /map="21q"
                        /clone="B137B7N13"

ORIGIN
Query Match      1.7%; Score 53; DB 10; Length 709;
Best Local Similarity 100.0%; Pred.No. 7.6e-14;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2888 TGAGCAGGTGATCACCCTGAGGCCAGAGGTTGAGACCAAGCTGCGCAACAT 2940
Db      393 TGAGCAGGTGATCACCCTGAGGCCAGAGGTTGAGACCAAGCTGCGCAACAT 445

RESULT 18
LOCUS      AG013687              709 bp    DNA          linear    GSS 16-FEB-2005
DEFINITION Homo sapiens genomic DNA, 21q region, clone: B137B7N13, genomic
survey sequence.
ACCESSION  AG013687
VERSION    AG013687.1  GI:3551635
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE  1
AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE     Homo sapiens genomic DNA, chromosome 21q
JOURNAL   Published Only in Database (1998)
REFERENCE  2 (bases 1 to 709)
AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE     Direct Submission
JOURNAL   Submitted (09-SEP-1998) Masahira Hattori, RIKEN Genomic Sciences
Center, RIKEN Yokohama Institute, Yokohama Research Promotion
Division, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa,
230-0045, Japan (E-mail:hattori@gscc.riken.jp, Tel:81-45-503-9111,
Fax:81-45-503-9113)
FEATURES   Location/Qualifiers
            source          1..709
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="21"
                        /map="21q"
                        /clone="B137B7N13"

ORIGIN
Query Match      1.7%; Score 53; DB 10; Length 709;
Best Local Similarity 100.0%; Pred.No. 7.6e-14;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2888 TGAGCAGGTGATCACCCTGAGGCCAGAGGTTGAGACCAAGCTGCGCAACAT 2940
Db      393 TGAGCAGGTGATCACCCTGAGGCCAGAGGTTGAGACCAAGCTGCGCAACAT 445

RESULT 19
LOCUS      AG013686              718 bp    DNA          linear    GSS 16-FEB-2005
DEFINITION Homo sapiens genomic DNA, 21q region, clone: B137B7N13, genomic
survey sequence.
ACCESSION  AG013686
VERSION    AG013686.1  GI:3551634
KEYWORDS   GSS.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE  1
AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE     Homo sapiens genomic DNA, chromosome 21q
JOURNAL   Published Only in Database (1998)
REFERENCE  2 (bases 1 to 718)
AUTHORS   Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
TITLE     Direct Submission
JOURNAL   Submitted (09-SEP-1998) Masahira Hattori, RIKEN Genomic Sciences
Center, RIKEN Yokohama Institute, Yokohama Research Promotion
Division, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa,
230-0045, Japan (E-mail:hattori@gscc.riken.jp, Tel:81-45-503-9111,
Fax:81-45-503-9113)
FEATURES   Location/Qualifiers
            source          1..718
                        /organism="Homo sapiens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:9606"
                        /chromosome="21"
                        /map="21q"
                        /clone="B137B7N13"

ORIGIN
Query Match      1.7%; Score 53; DB 10; Length 718;
Best Local Similarity 100.0%; Pred.No. 7.6e-14;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2888 TGAGCAGGTGATCACCCTGAGGCCAGAGGTTGAGACCAAGCTGCGCAACAT 2940
Db      396 TGAGCAGGTGATCACCCTGAGGCCAGAGGTTGAGACCAAGCTGCGCAACAT 448

RESULT 20
LOCUS      AA679936              354 bp    mRNA          linear    EST 02-DEC-1997
DEFINITION AC85G01.g1 Stratagene ovary (#937217) Homo sapiens cDNA clone
IMAGE:869424.3, similar to contains Alu repetitive element;contains
element TR1 repetitive element ;, mRNA sequence.
ACCESSION  AA679936
VERSION    AA679936.1  GI:2656403
KEYWORDS   EST.
SOURCE     Homo sapiens (human)
ORGANISM   Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.

REFERENCE  1 (bases 1 to 354)

```





## ORIGIN

Query Match 1.7%; Score 52; DB 1; Length 441;  
Best Local Similarity 100.0%; Pred. No. 2.4e-13;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3071 CAAGATTGTGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 3122  
|||||  
DB 217 CAAGATTGTGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 166

## RESULT 23

LOCUS CN277748 520 bp mRNA linear EST 16-MAY-2004  
DEFINITION 17000600183571 GRN\_PREHER Homo sapiens cDNA 5', mRNA sequence.  
ACCESSION CN277748  
VERSION CN277748.1 GI:47294162  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 520)  
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Muraue, J., Fisk, G.J.,  
Li, Y., Xu, C., Fang, R., Guegler, K., Rao, M.S., Mandalam, R.,  
Lebkowski, J. and Stanton, L.W.  
TITLE Transcriptome characterization elucidates signaling networks that  
control human ES cell growth and differentiation  
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)  
PUBMED 15146197  
COMMENT Contact: Brandenberger R  
Regenerative Medicine  
Geron Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@geron.com  
Insert Length: 520 Std Error: 0.00.  
Location/Qualifiers

FEATURES  
source 1..520  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cells, DMSO-treated H9 cell  
line"  
/clone\_1lb="GRN\_PREHER"  
/note="oligo dt primed, full-length enriched cDNA library  
from DMSO-treated hES cell line H9 (p22) maintained in  
feeder-free conditions"

## ORIGIN

Query Match 1.7%; Score 52; DB 7; Length 520;  
Best Local Similarity 100.0%; Pred. No. 2.4e-13;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3071 CAAGATTGTGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 3122  
|||||  
DB 467 CAAGATTGTGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 518

RESULT 24  
LOCUS AA680243/c 529 bp mRNA linear EST 02-DEC-1997  
DEFINITION ac86a01.s1 Stratsene ovary (#937217) Homo sapiens cDNA clone  
IMAGE:869448 3' similar to contains Alu repetitive element; contains  
element 11 repetitive element ;, mRNA sequence.  
ACCESSION AA680243  
VERSION AA680243.1 GI:2656211  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Homo sapiens

REFERENCE 1 (bases 1 to 529)  
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisels, G., Jost, S.,  
Kritzman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M.,  
Martin, J., Moore, B., Schellberg, K., Steptoe, M., Tan, F.,  
Thelshing, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.  
TITLE WashU-NCI human EST Project  
JOURNAL Unpublished (1997)  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
This clone is available royalty-free through LANT; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: -40m13 fwd. RT from Amersham  
High quality sequence stop: 359.  
Location/Qualifiers

FEATURES  
source 1..529  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:869448"  
/sex="female"  
/dev\_stage="49 year old"  
/lab\_host="SOLR cells (kanamycin resistant)"  
/clone\_1lb="Stratsene ovary (#937217)"  
/note="Organ: ovary; Vector: Bluescript SK; Site 1: EcoRI;  
Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt.  
Total ovary tissue, normal, caucasian. Average insert  
size: 0.8 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5'  
CAATCGGACGAG 3' -3' adaptor sequence: 5'  
CTCAGATTGTTTTTTTTTTT 3'

## ORIGIN

Query Match 1.7%; Score 52; DB 1; Length 529;  
Best Local Similarity 100.0%; Pred. No. 2.4e-13;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3071 CAAGATTGTGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 3122  
|||||  
DB 243 CAAGATTGTGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 192

RESULT 25  
LOCUS CK825142/c 553 bp mRNA linear EST 11-MAR-2004  
DEFINITION 1g44h08.y5 HR85 181et Homo sapiens cDNA clone IMAGE:5594343 5',  
mRNA sequence.  
ACCESSION CK825142  
VERSION CK825142.1 GI:44842067  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
1 (bases 1 to 553)  
Melton, D., Meadows, A., Clifton, S., Hillier, L., Marra, M., Pape, D.,  
Wylie, T., Martin, J., Bistacini, A., Schmitt, A., Thelshing, B.,  
Ritter, B., Ronko, I., Bennett, V., Cardenas, M., Gibbons, M.,  
McCann, R., Cole, R., Tsagaris, V., Williams, T., Jackson, Y. and  
Bowers, Y.  
TITLE WashU-Harvard Pancreas EST Project  
JOURNAL Unpublished (2000)  
COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812

Fax: 617-495-8557  
Email: dmetlon@biohp.harvard.edu  
This read is a 5' SEQUENCE of a previously sequenced pancreas clone  
Good hit to opposite strand read. . wrong orientation BUT PASSED FOR  
MOUSE-PANCREAS VERIFICATION  
Seq primer: -40UP from Gibco  
High quality sequence stop: 553.  
Location/Qualifiers

source  
1..555  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5594343"  
/issue\_type="Purified pancreatic islet"  
/lab\_host="DH10B"  
/clone\_id="HR85 islet"  
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site\_1:  
NotI; Site\_2: XhoI; cDNA made by oligo-dT priming.  
Size-selected on agarose gel. Average insert size ~1kb. 5'  
XhoI site was destroyed after directional cloning.  
Amplified once. Contact information: Hiroshi Inoue, MD,  
Metabolism Div. (Alan Permut Lab), Washington University  
School of Medicine, Box 8127, 660 South Euclid Ave., St.  
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:  
314-362-1916, Fax: 314-747-2692."

ORIGIN  
Query Match 1.7%; Score 52; DB 7; Length 553;  
Best Local Similarity 100.0%; Pred. No. 2.4e-13;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
3071 CAAGATTGGCCACTGCCTCCAGCTGGGCAACAGACAACTCTCTC 3122  
|||||  
540 CAAGATTGGCCACTGCCTCCAGCTGGGCAACAGACAACTCTCTC 489

ORIGIN  
Query Match 1.7%; Score 52; DB 7; Length 555;  
Best Local Similarity 100.0%; Pred. No. 2.3e-13;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
3071 CAAGATTGGCCACTGCCTCCAGCTGGGCAACAGACAACTCTCTC 3122  
|||||  
440 CAAGATTGGCCACTGCCTCCAGCTGGGCAACAGACAACTCTCTC 491

RESULT 26  
LOCUS CK825141 555 bp mRNA linear EST 11-MAR-2004  
DEFINITION 1944h08.x5 HR85 islet Homo sapiens cDNA clone IMAGE:5594343 3',  
mRNA sequence.  
CK825141  
ACCESSION CK825141  
VERSION CK825141.1 GI:44842066  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominae; Homo.  
1 (bases 1 to 555)  
Melton, D., Meadows, A., Clifton, S., Hillier, L., Marra, M., Pape, D.,  
Wyllie, T., Martin, J., Birstein, A., Schmitt, A., Theising, B.,  
Ritter, E., Ronko, I., Bennett, J., Cardenas, M., Gibbons, M.,  
McCann, R., Cole, R., Tsagaris, R., Williams, T., Jackson, Y. and  
Bowers, Y.  
WashU-Harvard Pancreas EST Project  
Unpublished (2000)  
Other ESTs: 1944h08.y1  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmetlon@biohp.harvard.edu  
This read is a 3' SEQUENCE of a previously sequenced pancreas  
clone  
Good hit to opposite strand read. . wrong orientation BUT PASSED FOR  
MOUSE-PANCREAS VERIFICATION  
Seq primer: -40RP from Gibco.  
Location/Qualifiers

RESULT 27  
LOCUS A0526715 568 bp DNA linear GSS 11-MAY-1999  
DEFINITION HS 5216 A2 C08 77A RPI-11 Human Male BAC Library Homo sapiens  
genomic clone Plate=792 Col=16 Row=E, genomic survey sequence.  
A0526715  
ACCESSION A0526715  
VERSION A0526715.1 GI:4769865  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominae; Homo.  
1 (bases 1 to 568)  
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,  
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and  
Hood, L.  
Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
10449764  
Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones are derived from the human BAC library RPI-11. For BAC  
library availability, please contact Peter de Jong  
(peter@dejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering\_bac.htm)  
or from Research Genetics (info@resgen.com). BAC end web server:  
http://www.hsc.washington.edu  
Plate: 792 row: E column: 16  
Seq primer: 17  
Class: BAC ends  
High quality sequence stop: 568.  
Location/Qualifiers  
1..568  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="Plate=792 Col=16 Row=E"  
/sex="male"

FEATURES  
source

1..568  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="Plate=792 Col=16 Row=E"  
/sex="male"

ORIGIN

/clone\_1lb="RPCT-11 Human Male BAC Library"  
/note="Vector: PBACE3.6; Site 1: EcoRI; Site 2: EcoRI;  
Male blood DNA was isolated from one randomly chosen donor  
and partially digested with a combination of EcoRI and  
EcoRI Methylase. Size selected DNA was cloned into the  
pBACE3.6 vector at EcoRI sites"

Query Match 1.7%; Score 52; DB 9; Length 568;  
Best Local Similarity 100.0%; Pred.No. 2.3e-13;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGAGGTGATGATCACTGAGGCGAGAGTTCGAGACCGCTGCGCAACAT 2940  
154 GAGGAGGTGATGATCACTGAGGCGAGAGTTCGAGACCGCTGCGCAACAT 205

RESULT 28  
B2606000 373 bp DNA linear GSS 08-JUN-2003  
LOCUS B2606000/c  
DEFINITION WHAS80TR Human MCF7 breast cancer cell line library (MCF7\_1) Homo  
sapiens genomic clone MCF7\_1-SN15, genomic survey sequence.  
ACCESSION B2606000  
VERSION B2606000.1 GI:31514561  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homniidae; Homo.  
REFERENCE 1 (bases 1 to 373)  
Volk,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,  
Gray,J.M. and Collins,C.  
End-sequence profiling: Sequence-based analysis of aberrant genomes  
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)  
12788976  
COMMENT Contact: Volk SV  
Colin Collins' lab  
UCSF Comprehensive Cancer Center  
UCSF Box 0808, San Francisco, CA 94143-0808, USA  
Tel: 415 502 7066  
Fax: 415 502 5665  
Email: svolik@cc.ucsf.edu  
This clone is available from Amplicon Express  
http://www.genomex.com  
Classes: BAC ends.

## FEATURES

## source

1. 373  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="MCF7\_1-SN15"  
/sex="female"  
/clone\_1lb="Human MCF7 breast cancer cell line library  
(MCF7\_1)"  
/note="Vector: pBACE3.6; Site 1: HindIII; This library was  
constructed from MCF7 breast cancer cell line by Amplicon  
Express (http://www.genomex.com) using their standard  
procedure."

## ORIGIN

Query Match 1.6%; Score 51; DB 9; Length 373;  
Best Local Similarity 100.0%; Pred.No. 7.3e-13;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGCGCACTGCACTCCAGCTGGGCAAGAGCAAGACTCTGTCT 3121  
185 CAAGATTGCGCACTGCACTCCAGCTGGGCAAGAGCAAGACTCTGTCT 135

RESULT 29  
CR545224

LOCUS CR545224 350 bp mRNA linear EST 07-JUL-2004  
DEFINITION DKFP459E1025\_r1\_459 (synonym: pcor1) Pongo pygmaeus cDNA clone  
DKFP459E1025\_5', mRNA sequence.  
ACCESSION CR545224  
VERSION CR545224.1 GI:49897049  
KEYWORDS EST.  
SOURCE Pongo pygmaeus (orangutan)  
ORGANISM Pongo pygmaeus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homniidae; Pongo.  
REFERENCE 1 (bases 1 to 350)  
Koehler,K., Beyer,A., Mewes,H.W., Well,B., Amlid,C., Oeanger,A.,  
Fobo,G., Han,M. and Wiemann,S.  
Pongo pygmaeus mRNA (Koehler,K., Beyer,A., Mewes,H.W., et al.)  
Unpublished (2004)  
COMMENT Contact: MIPS

MIPS  
Ingolstaedter Landster, J., D-85764 Neuberg, Germany  
This is the 5' sequence of the clone insert from S. Wiemann,  
Molecular Genome Analysis, German Cancer Research Center (DKFZ);  
Email: s.wiemann@dkfz-heidelberg.de; sequenced by BMFZ (Biomedical  
Research Center at the Heinrich-Heine-University,  
Duesseldorf/Germany) within the cDNA sequencing consortium of the  
German Genome Project. This clone (DKFP459E1025) is available at  
the RZPD in Berlin. Please contact the RZPD, Reagenzienzentrum,  
Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email:  
clone@rzpd.de Further information about the clone and the  
sequencing project is available at  
http://mips.gsf.de/projects/cdna/.

## FEATURES

## source

1. 350  
/organism="Pongo pygmaeus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9600"  
/clone="DKFP459E1025"  
/tissue\_type="cortex"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_1lb="459 (synonym: pcor1)"  
/note="Vector: pSport1\_sfi; Site\_1: SfiI; Site\_2: SfiIb"

## ORIGIN

Query Match 1.6%; Score 50; DB 7; Length 350;  
Best Local Similarity 100.0%; Pred.No. 2.2e-12;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCGCACTGCACTCCAGCTGGGCAAGAGCAAGACTCTGTCTC 3122  
98 AGATTGTGCGCACTGCACTCCAGCTGGGCAAGAGCAAGACTCTGTCTC 147

## RESULT 30

CB140816 471 bp mRNA linear EST 29-JAN-2003  
LOCUS K-EST0194308 L15CKK1 Homo sapiens cDNA clone L15CKK1-33-C11 5',  
mRNA sequence.  
DEFINITION CB140816  
ACCESSION CB140816  
VERSION CB140816.1 GI:28115050  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homniidae; Homo.

REFERENCE 1 (bases 1 to 471)  
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.

TITLE 21C Frontier Korean EST Project 2001  
JOURNAL Unpublished (2002)  
CONTACT: Kim YS  
COMMENT Genome Research Center

Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 33 row: C column: 11  
High quality sequence scop: 471.

## FEATURES

source

Location/Qualifiers  
1..471

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="L15CKX1-33-Cl1"  
/sex="M"  
/cell\_line="CK-K1"  
/lab\_host="Top10P"  
/clone\_1ib="L15CKX1"  
/note="Organ: Liver; Vector: PCNS-D2; Site 1: EcoRI;  
Site 2: NotI; The poly (A) + RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then deacapped  
with tobacco acid pyrophosphatase (TAP). The deacapped  
intact mRNA was ligated with DNA-RNA linker including  
EcoRI site by treatment of T4 RNA ligase and the first  
strand cDNA was synthesized from oligo dt-selected mRNA by  
priming with dt-tailed vector. The dt-tailed vector was  
adjusted to have about 60nt. The cDNA vector was  
circularized with B. coli DNA ligase after digestion of  
EcoRI which site is also included in vector. An RNA strand  
converted to a DNA strand by Okayama-Berg method. The  
obtained cDNA vectors were used for transformation of  
competent cells B. coli Top10P by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library."

## ORIGIN

Query Match 1.6%; Score 50; DB 6; Length 471;  
Best Local Similarity 100.0%; Pred. No. 2.2e-12;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2973 AATTGGCCGGCATGTGGCGCATGTGCTGCCAGCTACTCGGAGG 3022

Db 145 AATTGGCCGGCATGTGGCGCATGTGCTGCCAGCTACTCGGAGG 194

RESULT 31 A0487295 572 bp DNA linear GSS 24-APR-1999  
A0487295/c

LOCUS RPCI-11-245E13.TV RPCI-11 Homo sapiens genomic clone

DEFINITION RPCI-11-245E13, genomic survey sequence.

ACCESSION A0487295.1 GI:4673169

VERSION GSS.

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 572)

AUTHORS Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and

Venter, J.C.

TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready

Map Building

Unpublished (1997)

Other\_GSSs: RPCI-11-245E13.TV

Contact: Shaying Zhao, William Nierman, Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850

Tel: 301 838 0200

Fax: 301 838 0208

Email: hbeet1gr.org

Clones are derived from the human BAC library RPCI-11. For BAC

library availability, please contact Pieter de Jong

(pieter@delong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from  
Research Genet. ca ([http://www.tigr.org/tcdb/hungen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tcdb/hungen/bac_end_search/bac_end_search.html)).  
Seq primer: 17  
Class: BAC ends.

## FEATURES

source

Location/Qualifiers  
1..572

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7593804"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-245E13"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_1ib="RPCI-11"  
/note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;  
RPCI11 Human Male BAC Library"

## ORIGIN

Query Match 1.6%; Score 50; DB 9; Length 572;  
Best Local Similarity 100.0%; Pred. No. 2.2e-12;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGTCACATGCACTCCAGCCCTGGGCAAGAGCAAGACTGTCTC 3122

Db 572 AGATTGTCACATGCACTCCAGCCCTGGGCAAGAGCAAGACTGTCTC 523

RESULT 32 B47416/c  
LOCUS CIT978SK-161J21.TR CIT978SK Homo sapiens genomic clone 161J21,  
DEFINITION genomic survey sequence.

ACCESSION B47416

VERSION B47416.1 GI:2599822

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 634)

AUTHORS Adams, M.D., Rounsley, S.D., Field, C.B., Bass, S., Linher, K.,

Shizuya, H., Simon, M., and Venter, J.C.

Use of a human BAC End Sequence Database for Sequence-Ready Map

Building

Unpublished (1997)

Other\_GSSs: CIT978SK-161J21.TVB

Contact: Mark Adams

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: mdamas@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC

end search page:

[http://www.tigr.org/tcdb/hungen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tcdb/hungen/bac_end_search/bac_end_search.html)

Seq primer: M13 Reverse

Class: BAC ends.

## FEATURES

source

Location/Qualifiers  
1..634

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:5283048"  
/db\_xref="taxon:9606"  
/clone="161J21"  
/sex="Female"  
/cell\_type="Fibroblast"  
/clone\_1ib="CIT978SK"  
/note="Vector: pBAC108L; Site 1: HindIII; Site 2: HindIII;

ORIGIN CalTech Human BAC Library A"

Query Match 1.6%; Score 50; DB 9; Length 634;  
Best Local Similarity 100.0%; Pred. No. 2.2e-12;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGTGCACCTGCATCTCAGCTGGGCAACAGACAGACTCTGTCTC 3122  
|||||  
Db 586 AGATTGTGCACCTGCATCTCAGCTGGGCAACAGACAGACTCTGTCTC 537  
|||||

RESULT 33  
AG145253 656 bp DNA linear GSS 08-JAN-2002  
LOCUS Pan troglodytes DNA, clone: RP43-006H23.T7, genomic survey  
DEFINITION sequence.  
ACCESSION AG145253  
VERSION AG145253.1 GI:16674931  
KEYWORDS GSS:  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Pan.

REFERENCE 1  
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,  
Totoki, Y., Watanabe, H. and Sakaki, Y.  
TITLE BAC end sequences of library RP43-43  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 656)  
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,  
Totoki, Y., Watanabe, H. and Sakaki, Y.  
TITLE Direct Submission  
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),  
1-7-22 Suenho-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail: schimpes@sc.riken.go.jp, URL: http://hgc.gsc.riken.go.jp/  
Tel: 81-45-503-9111, Fax: 81-45-503-9170)  
COMMENT Clones are derived from the chimpanzee BAC library RP43-43 This BAC  
end was generated during the Rad process and may have higher chance  
of clone tracking errors.  
PRIMERS  
Sequencing: T7  
LIBRARY  
Vector : pBACe3.6  
R.Site 1 : EcoRI  
R.Site 2 : EcoRI.  
Location/Qualifiers  
FEATURES  
source  
1..656  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone="RP43-006H23.T7"  
/sex="male"  
/cell\_type="lymphocytes"  
/clone\_lib="RP43-43 Chimpanzee Male BAC Library"

ORIGIN

Query Match 1.6%; Score 50; DB 10; Length 656;  
Best Local Similarity 100.0%; Pred. No. 2.2e-12;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGTGCACCTGCATCTCAGCTGGGCAACAGACAGACTCTGTCTC 3122  
|||||  
Db 394 AGATTGTGCACCTGCATCTCAGCTGGGCAACAGACAGACTCTGTCTC 443  
|||||

RESULT 34  
AV734315 694 bp mRNA linear EST 17-OCT-2000  
LOCUS AV734315 cda Homo sapiens cDNA clone cdANCO5 5', mRNA sequence.  
DEFINITION  
ACCESSION AV734315

VERSION AV734315.1 GI:10851860  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

REFERENCE 1 (bases 1 to 694)  
AUTHORS Yang, Y., Song, H., Peng, Y., Gu, Y., Gao, G., Xiao, H., Xu, X., Li, N.,  
Qian, B., Liu, F., Qu, J., Gao, K., Cheng, Z., Xu, Z., Zeng, L., Xu, S.,  
Qi, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R.,  
Chen, J., Chen, Z. and Han, Z.  
TITLE Homo sapiens cDNA clones  
JOURNAL Unpublished (2000)  
COMMENT Contact: Zeguang Han  
Chinese National Human Genome Center at Shanghai  
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai  
201203, P. R. China  
Tel: 86-21-50801919 (ex. 45)  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.  
Location/Qualifiers  
FEATURES  
source  
1..694  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="cdANCO5"  
/tissue\_type="pheochromocytoma"  
/dev\_stage="Adult"  
/lab\_host="BM25.8"  
/clone\_lib="cda"  
/note="Vector: pTriplEx2; Site\_1: sf1A; Site\_2: sf1B"

ORIGIN

Query Match 1.6%; Score 50; DB 1; Length 694;  
Best Local Similarity 100.0%; Pred. No. 2.2e-12;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGTGCACCTGCATCTCAGCTGGGCAACAGACAGACTCTGTCTC 3122  
|||||  
Db 397 AGATTGTGCACCTGCATCTCAGCTGGGCAACAGACAGACTCTGTCTC 446  
|||||

RESULT 35  
CN413517 706 bp mRNA linear EST 16-MAY-2004  
LOCUS CN413517  
DEFINITION 17000532630078 GRN\_EB Homo sapiens cDNA 5', mRNA sequence.  
ACCESSION CN413517  
VERSION CN413517.1 GI:47401111  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

REFERENCE 1 (bases 1 to 706)  
AUTHORS Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G. J.,  
Li, Y., Xu, C., Fang, R., Guejler, K., Rao, M. S., Mandalam, R.,  
Lebkoweki, J. and Stanton, L. W.  
TITLE Transcriptional characterization elucidates signaling networks that  
control human ES cell growth and differentiation  
JOURNAL Nat. Biotechnol. 22 (6), 707-716 (2004)  
PUBMED 15146197  
COMMENT Contact: Brandenberger R  
Regenerative Medicine  
Geron Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@geron.com  
Insert length: 706 Std Error: 0.00.  
Location/Qualifiers

source

1..706  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="embryonic stem cells", embryo bodies  
derived from H1, H7 and H9 cells"  
/clone\_1lb="GRN BP"  
/note="oligo dt primed, full-length enriched cDNA library  
from embryo body outgrowths derived from hES cell lines  
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free  
conditions."

ORIGIN

Query Match 1.6%; Score 50; DB 7; Length 706;  
Best Local Similarity 100.0%; Pred. No. 2.2e-12;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGTGCACTGCATCTCCAGCTGGGCAAGAGCAAGACTGTCTC 3122  
|||||  
Db 134 AGATTGTGCACTGCATCTCCAGCTGGGCAAGAGCAAGACTGTCTC 183

RESULT 36  
CD357076 770 bp mRNA linear EST 29-MAY-2003  
LOCUS IMAGE:30401040 5', mRNA sequence.  
DEFINITION  
ACCESSION CD357076  
VERSION CD357076.1 GI:31128487  
KEYWORDS EST.  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 770)  
NIH-MGC http://mgi.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgaiba-remail.nih.gov  
Tissue Procurement: Dr. Michael Brownstein and Dr. Miklos Palkovits  
cDNA Library Preparation: CLOUTEC Laboratories, Inc.  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/MLNL at:  
http://image.llnl.gov  
Plate: NDCM181 row: C column: 01  
High quality sequence ecop: 538.  
Location/Qualifiers  
1..770  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_1lb="IMAGE:30401040"  
/lab\_host="DH10B (T1 phage-resistant)"  
/note="Organ: Blood vessels - aorta, basilar and artery;  
vector: pDNR-LIB, Site 1: SfiI (ggcgcatggcc); Site 2:  
SfiI (ggcgctcgcc); 5' and 3' adaptors were used in  
cloning as follows: 5' adaptor sequence:  
5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence:  
5'-ATTCTAGAGCGCGCGCCGACATG-DT(30)BN-3' (where B = A,  
C, or G and N = A, C, G, or T). Average insert size 1.4 kb  
(range 0.5-4.0 kb). 14/15 colonies contained inserts by  
PCR. This library was enriched for full-length clones and  
was constructed by Clontech Laboratories (Palo Alto, CA).  
Note: this is a NIH\_MGC library."

FEATURES  
source

Query Match 1.6%; Score 50; DB 6; Length 770;  
Best Local Similarity 100.0%; Pred. No. 2.2e-12;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGTTGTGCACTGCATCTCCAGCTGGGCAAGAGCAAGACTGTCTC 3120  
|||||  
Db 382 CAAGTTGTGCACTGCATCTCCAGCTGGGCAAGAGCAAGACTGTCTC 431

RESULT 37  
AY758761 797 bp DNA linear GSS 14-APR-2005  
LOCUS AY758761  
DEFINITION RP43-99b TV Pan troglodytes genomic DNA Pan troglodytes genomic  
clone RP43-99b TV, genomic survey sequence.  
ACCESSION AY758761  
VERSION AY758761.1 GI:62544847  
KEYWORDS GSS.  
SOURCE  
ORGANISM Pan troglodytes (chimpanzee)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Pan.  
1 (bases 1 to 797)  
Yohn,C.T., Jiang,Z., McGrath,S.D., Hayden,K.E., Khalovich,P.,  
Johnson,M.E., Eichler,M.Y., McPherson,J.D., Zhao,S., Pabdo,S. and  
Richler,E.B.  
Lineage-specific expansions of retroviral insertions within the  
genomes of African great apes but not humans and orangutans  
PLoS Biol. 3 (4), e110 (2005)  
15737067  
Contact: Evan Eichler  
Department of Genetics  
Case Western Reserve University  
2109 Adelbert Rd., BRB 720, Cleveland, OH 44106, USA  
Tel: 216 368 4883  
Fax: 216 368 3432  
Email: eee@po.cwru.edu  
TJ forward and TV reverse  
Clas: BAC ends.  
Location/Qualifiers  
1..797  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone\_1lb="RP43-99b TV"  
/clone\_1lb="Pan troglodytes genomic DNA"

FEATURES  
source

ORIGIN

Query Match 1.6%; Score 49; DB 9; Length 797;  
Best Local Similarity 100.0%; Pred. No. 6.7e-12;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGTGCACTGCATCTCCAGCTGGGCAAGAGCAAGACTGTCT 3121  
|||||  
Db 404 AGATTGTGCACTGCATCTCCAGCTGGGCAAGAGCAAGACTGTCT 452

RESULT 38  
AM006453 452 bp mRNA linear EST 10-SEP-1999  
LOCUS AM006453/C  
DEFINITION w16h01.x1 NCI CGAP P122 Homo sapiens cDNA clone IMAGE:2481745 3'  
similar to contains Alu repetitive element", mRNA sequence.  
ACCESSION AM006453  
VERSION AM006453.1 GI:5855231  
KEYWORDS EST.  
SOURCE  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 452)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.

REFERENCE  
AUTHORS



**TITLE** National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
**JOURNAL** Unpublished (1997)  
**COMMENT** Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-r@mail.nih.gov  
 Tissue Procurement: Michael J. Brownstein, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/BLNI at: [www-bio.ihl.nih.gov/bbrp/image/image.html](http://www-bio.ihl.nih.gov/bbrp/image/image.html)  
 Seq primer: -40up from G1hco  
 High quality sequence stop: 423.  
**FEATURES**  
 source  
 1..452  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:2481745"  
 /sex="male"  
 /tissue\_type="normal prostate"  
 /lab\_host="DH10B"  
 /clone\_1ib="NCI CGAP Pr22"  
 /note="Organ: prostate; Vector: pT7T3D-Pac (pharmacia) with a modified polylinker; 1st strand cDNA was prepared from normal prostate bulk tissue, and was then primed with a Not I - o11go(dT) primer. Double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library is normalized, and was constructed by Bento Soares and M. Patricia Bonaldo."

**ORIGIN**  
 Query Match 1.5%; Score 48; DB 1; Length 452;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;  
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 3073 AGATTGTGCGACTGCCTCCAGCTGGGCAACAGACGACTGTCTC 3120  
 Db 73 AGATTGTGCGACTGCCTCCAGCTGGGCAACAGACGACTGTCTC 26

**RESULT 39**  
 LOCUS AQ757333 548 bp DNA linear GSS 27-JUL-1999  
 DEFINITION HS 5484 AI H10 T7A RPCI-11 Human Male BAC Library Homo sapiens genomic clone Plate=1060 Col=19 Row=0, genomic survey sequence.  
 ACCESSION AQ757333  
 VERSION AQ757333.1 GI:5622184  
 KEYWORDS GSS.  
 ORGANISM Homo sapiens (human)  
 SOURCE Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo  
 1 (bases 1 to 548)  
 Mahitras,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.  
**REFERENCE**  
 AUTHORS 1 (bases 1 to 548)  
 Mahitras,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.  
**TITLE** Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome  
**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
**PUBMED** 10449764  
**COMMENT** Contact: Mahitras GG, Wallace JC, Hood L  
 High Throughput Sequencing Center  
 University of Washington  
 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887  
 Email: jwallace@u.washington.edu  
 Clones are derived from the human BAC library RPCI-11. For BAC

library availability, please contact Pieter de Jong ([pieter@jlong.med.buffalo.edu](mailto:pieter@jlong.med.buffalo.edu)). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering/bac.htm>) or from Research Genetics (<http://www.htsc.washington.edu>)  
 Plate: 1060 row: 0 column: 19  
 Seq primer: 77  
 Class: BAC ends  
 High quality sequence stop: 548.  
**FEATURES**  
 source  
 1..548  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="Plate=1060 Col=19 Row=0"  
 /sex="male"  
 /clone\_1ib="RPCI-11 Human Male BAC Library"  
 /note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI; male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI methylase. Size selected DNA was cloned into the pBACe3.6 vector at EcoRI sites"

**ORIGIN**  
 Query Match 1.5%; Score 48; DB 9; Length 548;  
 Best Local Similarity 100.0%; Pred. No. 2.1e-11;  
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 3075 ATTGTGCGACTGCCTCCAGCTGGGCAACAGACGACTGTCTC 3122  
 Db 446 ATTGTGCGACTGCCTCCAGCTGGGCAACAGACGACTGTCTC 399

**RESULT 40**  
 LOCUS AQ121008/c 457 bp DNA linear GSS 22-SEP-1998  
 DEFINITION HS 3072 A2 C01 MF CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=3073 Col=2 Row=E, genomic survey sequence.  
 ACCESSION AQ121008  
 VERSION AQ121008.1 GI:3498174  
 KEYWORDS GSS.  
 ORGANISM Homo sapiens (human)  
 SOURCE Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 457)  
 Mahitras,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.  
**REFERENCE**  
 AUTHORS 1 (bases 1 to 457)  
 Mahitras,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.  
**TITLE** Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome  
**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
**PUBMED** 10449764  
**COMMENT** Contact: Mahitras GG, Wallace JC, Hood L  
 High Throughput Sequencing Center  
 University of Washington  
 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887  
 Email: jwallace@u.washington.edu  
 Sequence Tagged Connector  
 Plate: 3073 row: E column: 2  
 Class: BAC ends  
 High quality sequence stop: 457.  
**FEATURES**  
 source  
 1..457  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="Plate=3073 Col=2 Row=E"  
 /sex="male"



COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: David N. Louis, M.D.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>  
Plate: LLM9786 row: 0 column: 01  
High quality sequence stop: 734.  
Location/Qualifiers  
1..847

FEATURES  
source  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4213848"  
/tissue\_type="glioblastoma with EGFR amplification"  
/lab\_host="DH10B (71 phage-resistant)"  
/clone\_1ib="NCI CGAP Brn64"  
/note="Organ: Brain; Vector: pCMV-SPORT6; Site\_1: NotI; Site\_2: SalI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.57 kb. Constructed by Life Technologies. Note: this is a NCI CGAP library."  
Technology: Note: this is a NCI CGAP library."

ORIGIN

Query Match 1.5%; Score 47; DB 2; Length 847;  
Best Local Similarity 100.0%; Pred. No. 6.4e-11;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2888 TGAGCAGGTGATCACCCTGAGCCGAGGAGTTGAGACAGCCTGGC 2934  
|||||  
Db 128 TGAGCAGGTGATCACCCTGAGCCGAGGAGTTGAGACAGCCTGGC 82

RESULT 44  
AA078268/c 216 bp mRNA linear EST 24-SEP-1999  
LOCUS 7H48E05 Chromosome 7 HeLa cDNA Library Homo sapiens cDNA clone  
DEFINITION 7H48E05, mRNA sequence.  
ACCESSION AA078268  
VERSION AA078268.1 GI:1837742  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 216)  
Touchman, J.W., Bouffard, G.G., Weintraub, L.A., Idol, J.R., Wang, L., Robbins, C.M., Nusbaum, J.C., Lovett, M., and Green, E.D.  
2006 expressed-sequence tags derived from human chromosome 7-enriched cDNA libraries  
Genome Res. 7 (3), 281-292 (1997)  
9074931  
Contact: Eric D. Green  
Genome Technology Branch  
National Human Genome Research Institute/NIH  
49 Convent Dr., MSC4431, Building 49, Room 2A08, Bethesda, MD 20892  
Tel: 3014020201  
Fax: 3014024735  
Email: [egreen@nhgri.nih.gov](mailto:egreen@nhgri.nih.gov)  
Plate: 48 row: 8 column: 05  
Seq primer: -21M13 (ABI).  
Location/Qualifiers  
1..216

FEATURES  
source  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="7H48E05"  
/sex="female"  
/cell\_line="HeLa cell line; ATCC"  
/lab\_host="E. coli strain DH5 alpha"

ORIGIN

Query Match 1.5%; Score 46; DB 1; Length 216;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2888 TGAGCAGGTGATCACCCTGAGCCGAGGAGTTGAGACAGCCTGG 2933  
|||||  
Db 82 TGAGCAGGTGATCACCCTGAGCCGAGGAGTTGAGACAGCCTGG 37

RESULT 45  
AI309360 255 bp mRNA linear EST 08-APR-1999  
LOCUS tbs1c03.x1 NCI CGAP Br15 Homo sapiens cDNA clone IMAGE:2057860 3'  
DEFINITION similar to contains Alu repetitive element; mRNA sequence.  
ACCESSION AI309360  
VERSION AI309360.1 GI:4004231  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 255)  
NCI CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgabbs-remail.nih.gov  
Tissue Procurement: Dennis Sgroi, M.D., Kristina Cole, M.D., Ph.D.  
Student, Michael R. Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: David B. Kitzman, Ph.D.  
cDNA Library Arrayed by: Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: [www.bio.llnl.gov/bbtp/image/image.html](http://www.bio.llnl.gov/bbtp/image/image.html)  
Insert Length: 317 Std Error: 0.00  
Seq primer: -40UP from Glibco.  
Location/Qualifiers  
1..255

FEATURES  
source  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2057860"  
/sex="female"  
/tissue\_type="adenocarcinoma"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_1ib="NCI CGAP Br15"  
/note="Organ: breast; Vector: PAMPI; mRNA made from breast adenocarcinoma tissue, cDNA made by oligo-dT priming. directionally cloned. size-selected on agarose gel, average insert size 400 bp. Primary library, non-amplified."

ORIGIN

Query Match 1.5%; Score 46; DB 1; Length 255;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2895 GGTGATCAGCTGAGGCCGAGGAGTTGAGACAGCCTGGCAACAT 2940  
|||||

Db 166 GGTGATCACCCTGAGGCCAGAGTTCCAGACCAAGCCTGGCCAAACAT 211

RESULT 46  
BX475355 313 bp mRNA linear EST 04-SEP-2003

LOCUS  
DEFINITION DKFZp666M14176.r1.686 (synonym: h1cc3) Homo sapiens cDNA clone

ACCESSION BX475355  
VERSION BX475355.1 GI:31672638

KEYWORDS  
SOURCE EST.

ORGANISM Homo sapiens (human)

REFERENCE  
AUTHORS Bhatia, A., Lauber, J., Mewes, H.W., Well, B., Amid, C., Oanger, A.,  
Fob, G., Han, M., and Wiemann, S.  
EST (Bhatia, A., Lauber, J., Mewes, H.W., Well, B., et al.)  
Unpublished (2003)

TITLE JOURNAL  
COMMENT

FEATURES  
source  
1. 313  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp666M14176"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_1b="686 (synonym: h1cc3)"  
/note="Vector: pTRIPLEX2; Site\_1: SfiI; Site\_2: SfiI;  
CDNA-collection"

ORIGIN  
Query Match 1.5%; Score 46; DB 5; Length 313;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACCCTGAGGCCAGAGTTCCAGACCAAGCCTGGCCAAACAT 2940  
|||||  
173 GGTGATCACCCTGAGGCCAGAGTTCCAGACCAAGCCTGGCCAAACAT 92

Db 47 GGTGATCACCCTGAGGCCAGAGTTCCAGACCAAGCCTGGCCAAACAT 92

RESULT 47  
AW841960 314 bp mRNA linear EST 18-MAY-2000

LOCUS  
DEFINITION IL5-CN0024-100300-030-H12 CN0024 Homo sapiens cDNA, mRNA sequence.

ACCESSION AW841960  
VERSION AW841960.1 GI:7935943

KEYWORDS  
SOURCE EST.

ORGANISM Homo sapiens (human)

REFERENCE  
AUTHORS Dias Neto, B., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V.,  
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

TITLE Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
Sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

JOURNAL  
PUBMED 10737800

COMMENT  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the PAPSP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=et2-IL5-CN0024-100  
300-030-H12&3=2000-03-10&4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 297.

FEATURES  
source  
1. 314  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_1b="CN0024"  
/note="Organ: colon, normal; Vector: puc18; Site\_1: SmaI;  
Site\_2: SmaI; A mini-library was made by cloning products  
derived from ORFESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

ORIGIN  
Query Match 1.5%; Score 46; DB 1; Length 314;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACCCTGAGGCCAGAGTTCCAGACCAAGCCTGGCCAAACAT 2940  
|||||  
173 GGTGATCACCCTGAGGCCAGAGTTCCAGACCAAGCCTGGCCAAACAT 218

Db 173 GGTGATCACCCTGAGGCCAGAGTTCCAGACCAAGCCTGGCCAAACAT 218

RESULT 48  
CD514740 319 bp mRNA linear EST 06-JUN-2003

LOCUS  
DEFINITION AGENCOURT\_14394439 NIH-MGC\_181 Homo sapiens cDNA clone

ACCESSION CD514740  
VERSION CD514740.1 GI:31446458

KEYWORDS  
SOURCE EST.

ORGANISM Homo sapiens (human)

REFERENCE  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE  
JOURNAL  
COMMENT  
Unpublished (1999)  
Contact: Daniela S. Gerhard, Ph.D.  
Office of Cancer Genomics  
National Cancer Institute / NIH  
Bldg. 31 Rm10A07 Bethesda, MD 20892  
Email: cgsabbs-remail.nih.gov  
Tissue Procurement: Dr. Michael Brownstein  
cDNA Library Preparation: Invitrogen Corp  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNLN at:  
http://image.llnl.gov

Plate: NDAM491 row: f column: 24  
High quality sequence stop: 319.  
Location/Qualifiers

## FEATURES

source

1. 319

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:30408431"

/issue\_type="White Matter"

/dev\_stage="Unknown"

/lab\_host="PH10B-Ton A ( T1 and T5 phage resistances) "

/clone\_1ib="NIH MGC 181"

/note="Vector: PCMV-SPORT6.1; Site 1: NotI; Site 2: EcoRV

(destroyed); library is oligo-dt primed and directionally

cloned (Gcovr site is destroyed upon cloning). Average

insert size 1.42 kb. Library was constructed by

(Invitrogen). Note: this is a NIH\_MGC library."

## ORIGIN

Query Match

Best Local Similarity 1.5%; Score 46; DB 6; Length 319;

Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2895 GGTGATCACCCTGAGCCAGAGTTCCAGACCAAGCTGGCCAAACAT 2940

Db 116 GGTGATCACCCTGAGCCAGAGTTCCAGACCAAGCTGGCCAAACAT 71

## RESULT 49

AA661680/c

DEFINITION nu87b06.s1 NCI CGAP Alvi1 Homo sapiens cDNA clone IMAGE:1217651

LOCUS AA661680 329 bp mRNA linear EST 12-NOV-1997

ACCESSION AA661680 similar to contains Alu repetitive element; mRNA sequence.

VERSION AA661680.1 GI:2615771

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homnidae; Homo.

1 (bases 1 to 329)

NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.

National Cancer Institute, Cancer Genome Anatomy Project (CGAP),

Tumor Gene Index

Unpublished (1997)

Contact: Robert Strausberg, Ph.D.

Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)

Tissue Procurement: Lee Helman, M.D., Michael R. Emmert-Buck, M.D.,

Ph.D.

CDNA library Preparation: David B. Kitzman, Ph.D.

DNA Sequencing by: Greg Lennon, Ph.D.

Clone distribution: NCI-CGAP clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

[www.bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)

Seg primer: -80m3 fwd. Rf from Amersham

High quality sequence stop: 309.

Location/Qualifiers

1. 329

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:1217651"

/issue\_type="Alveolar rhabdomyosarcoma"

/lab\_host="DH10B"

/clone\_1ib="NCI CGAP Alvi1"

/note="Vector: PAMPI1; mRNA made from alveolar

rhabdomyosarcoma, cDNA made by oligo-dt priming.

Non-directionally cloned. Size-selected on agarose gel,

average insert size 600 bp. Reference: Kitzman et al,

(1996) Cancer Research 56:5380-5383."

## ORIGIN

Query Match 1.5%; Score 46; DB 1; Length 329;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2895 GGTGATCACCCTGAGCCAGAGTTCCAGACCAAGCTGGCCAAACAT 2940

Db 241 GGTGATCACCCTGAGCCAGAGTTCCAGACCAAGCTGGCCAAACAT 196

## RESULT 50

BE061760/c

LOCUS BE061760 341 bp mRNA linear EST 09-JUN-2000

DEFINITION RCI-BT0254-111099-012-a10 BT0254 Homo sapiens cDNA, mRNA sequence.

ACCESSION BE061760

VERSION BE061760.1 GI:8406410

KEYWORDS EST.

SOURCE Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homnidae; Homo.

1 (bases 1 to 341)

Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,

Nagai, M.A., da Silva, M. Jr., Zago, M.A., Bordin, S., Costa, F.F.,

Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,

Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V.,

O'Hare, M.U., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

Simpson, A.J.

Shotgun sequencing of the human transcriptome with ORF expressed

sequence tags

Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

Contact: Simpson A.J.G.

Laboratory of Cancer Genetics

Ludwig Institute for Cancer Research

Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

Brazil

Tel: +55-11-2704922

Fax: +55-11-2707001

Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)

This sequence was derived from the PAPERB/LICR Human Cancer Genome

Project. This entry can be seen in the following URL

(<http://www.ludwig.org.br/scripts/gethtml2.pl?cl=tr2-RC1-BT0254-111>

099-012-a10&amp;t3=1999-10-11&amp;t4=1)

Seq primer: puc 18 forward

High quality sequence start: 83

High quality sequence stop: 139.

Location/Qualifiers

1. 341

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/dev\_stage="Adult"

/clone\_1ib="BT0254"

/note="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2:

SmaI; A mini-library was made by cloning products derived

from ORESTES PCR (U.S. Letters Patent application No.

196,716 - Ludwig Institute for Cancer Research) profiles

into the pUC 18 vector. Reverse transcription of tissue

mRNA and cDNA amplification were performed under low

stringency conditions."

## ORIGIN

Query Match

Best Local Similarity 1.5%; Score 46; DB 2; Length 341;

Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTCACCTGAGCCAGAGTTCCAGACCAAGCTGGCCAAACAT 3116

Db 166 CAAGATTGTCACCTGAGCCAGAGTTCCAGACCAAGCTGGCCAAACAT 121

## ORIGIN

RESULT 51  
BF6767274 372 bp mRNA linear EST 12-JAN-2001  
LOCUS RC2-CN0084-091200-011-cl12 CN0084 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF6767274  
VERSION BF6767274.1 GI:12115174  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
REFERENCE 1 (bases 1 to 372)  
Dias Neto, E., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V.,  
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
JOURNAL 10737800  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=RC2&t2=RC2-CN0084-  
091200-011-cl12&t3=2000-12-09&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 372.  
Location/Qualifiers  
1..372  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_1lb="CN0084"  
/notes="Organ: colon normal; Vector: puc18; Site\_1: Sma1;  
Site\_2: Sma1; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the puc 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

ORIGIN  
Query Match 1.5%; Score 46; DB 2; Length 372;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACTGAGCGCAGAGCTTGAGACCAAGCTGCGCAACAT 2340  
|||||  
181 GGTGATCACTGAGCGCAGAGCTTGAGACCAAGCTGCGCAACAT 226  
|||||

RESULT 52  
AQ201612/c 387 bp DNA linear GSS 20-APR-1999  
LOCUS RPCI11-59M4.TK RPCI-11 Homo sapiens genomic clone RPCI-11-59M4,  
DEFINITION genomic survey sequence.  
ACCESSION AQ201612  
VERSION AQ201612  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE 1 (bases 1 to 387)  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
AUTHORS Adams, M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,  
Bexry, K., Granger, D., Suh, E., Wible, C., de Jong, P. and Venter, J.C.  
TITLE Use of human BAC End Sequences for Sequence-Ready Map Building  
JOURNAL Unpublished (1998)  
COMMENT Other\_GSSs: RPCI11-59M4.TK  
Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdadams@tigr.org  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieter@jlong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
Research Genetics (info@resgen.com). BAC end search page:  
http://www.tigr.org/cdb/hungen/bac\_end\_search/bac\_end\_search.html  
Classes: BAC ends.  
Location/Qualifiers  
1..387  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7522563"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-59M4"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_1lb="RPCI-11"  
/notes="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
RPCI11 Human Male BAC Library"

ORIGIN  
Query Match 1.5%; Score 46; DB 9; Length 387;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAAGTTGTGCACCTGCACCTCCAGCTTGGCAACAGACGACATC 3116  
|||||  
289 CAAAGTTGTGCACCTGCACCTCCAGCTTGGCAACAGACGACATC 244  
|||||

RESULT 53  
AQ134220 389 bp DNA linear GSS 23-SEP-1998  
LOCUS AQ134220/c  
DEFINITION HS 3047 B1 A12 MF CIT Approved Human Genomic Sperm Library D Homo  
sapiens genomic clone Plate=3047 Col=23 Row=B, genomic survey  
sequence.  
ACCESSION AQ134220  
VERSION AQ134220.1 GI:3525586  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
REFERENCE 1 (bases 1 to 389)  
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,  
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and  
Hood, L.  
TITLE Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
PUBMED 10449764  
COMMENT Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887

Email: jwallace@u.washington.edu  
Sequence Tagged Connector  
Plate: 3047 row: B column: 23  
Class: BAC ends  
High quality sequence stop: 389.  
Location/Qualifiers

## FEATURES

source

1.389  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="Plate:3047 Col=23 Row=B"  
/sex="male"  
/clone\_1ib="CIT Approved Human Genomic Sperm Library D"  
/note="Organ: sperm; Vector: pBeloBAC11, BAC clones in  
E-Coli DH10B"

## ORIGIN

Query Match 1.5%; Score 46; DB 9; Length 389;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAGATTGTGCCACTGCACTCCAGCTCGGCAACAGAGCAAGACTC 3116  
DB 266 CAGATTGTGCCACTGCACTCCAGCTCGGCAACAGAGCAAGACTC 221

RESULT 54  
AA503019/c 405 bp mRNA linear EST 15-AUG-1997  
LOCUS ng20d04.s1 NCI CGAP OV2 Homo sapiens cDNA IMAGE:929959  
DEFINITION similar to contains Alu repetitive element; mRNA sequence.  
ACCESSION AA503019  
VERSION AA503019.1 GI:2237986  
KEYWORDS EST.

## SOURCE

Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homniidae; Homo.  
1 (bases 1 to 405)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)

## JOURNAL

COMMENT

Contact: Robert Strausberg, Ph.D.  
Email: cgapdb-r@mail.nih.gov  
Tissue Procurement: Christopher A. Moskaluk, M.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: David B. Kitzman, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/BLNT at:  
[www-bio.11nl.gov/db/rdp/image/image.html](http://www-bio.11nl.gov/db/rdp/image/image.html)  
Insert length: 1127 Std Error: 0.00  
Seq primer: -40m3 fwd. ET from Amerham  
High quality sequence stop: 392.  
Location/Qualifiers

## FEATURES

source

1.405  
/organism="Homo sapiens"  
/mol\_type="RNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:929959"  
/sex="female"  
/tissue\_type="ovary"  
/lab\_host="DH10B"  
/clone\_1ib="NCI CGAP OV2"  
/note="Vector: PAMPI0; mRNA made from invasive ovarian  
tumor, cDNA made by oligo-dt priming. Non-directionally  
cloned. Size-selected on agarose gel, average insert size  
600 bp. Reference: Kitzman et al. (1996) Cancer Research  
56:5380-5383."

## ORIGIN

Query Match 1.5%; Score 46; DB 1; Length 405;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAGATTGTGCCACTGCACTCCAGCTCGGCAACAGAGCAAGACTC 3116  
DB 106 CAGATTGTGCCACTGCACTCCAGCTCGGCAACAGAGCAAGACTC 61

## RESULT 55

AQ284177

LOCUS

DEFINITION RPEC11-80L20\_TV RPEC1-11 Homo sapiens genomic clone RPEC1-11-80L20,  
genomic survey sequence.

ACCESSION AQ284177  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homniidae; Homo.  
1 (bases 1 to 406)  
Adams M.D., Rounsley, S.D., Zhao, S., Bass, S., Linher, K., Golden, K.,  
Berry, K., Granger, D., Suh, E., Wible, C., de Jong, P., and Venter, J.C.  
Use of human BAC End Sequences for Sequence-Ready Map Building  
Unpublished (1998)  
Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdamad@tigr.org

Clones are derived from the human BAC library RPEC1-11. For BAC  
library availability, please contact Pieter de Jong  
([pieterdejong.med.buffalo.edu](mailto:pieterdejong.med.buffalo.edu)). Clones may be purchased from  
BACRAC Resources (<http://bacpac.med.buffalo.edu/ordering>) or from  
Research Genetics ([intoregen.com](http://intoregen.com)). BAC end search page:  
[http://www.tigr.org/cdb/hungun/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/cdb/hungun/bac_end_search/bac_end_search.html)  
Seq primer: 17  
Class: BAC ends.  
Location/Qualifiers

## FEATURES

source

1.406  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7530619"  
/db\_xref="taxon:9606"  
/clone="RPEC1-11-80L20"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_1ib="RPEC1-11"  
/note="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI;  
RPEC11 Human Male BAC Library"

## ORIGIN

Query Match 1.5%; Score 46; DB 9; Length 406;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATACCTGAGGCGAGGAGTTCGAGACCGCTGGCCAACAT 2940  
DB 25 GGTGATACCTGAGGCGAGGAGTTCGAGACCGCTGGCCAACAT 70

## RESULT 56

AW815323/c

LOCUS

DEFINITION QV0-ST0215-060100-083-d08 ST0215 Homo sapiens cDNA, mRNA sequence.  
ACCESSION AW815323  
VERSION AW815323.1 GI:7908317  
KEYWORDS EST.



**SOURCE** Homo sapiens (human)

**ORGANISM** Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.

**REFERENCE** Dias Neto, R., Garcia Correa, R., Verjowski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.P.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,  
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.

**TITLE** Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags

**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

**COMMENT** 10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the PAPER/PICR Human Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/ecrpts/gethtml2.pl?cl=kt2-QV0-ST0215-060  
100-083-d08kt3=2000-01-06kt4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 13  
High quality sequence stop: 445.  
Location/Qualifiers  
1..445  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_1lb="ST0215"  
/note="Organ: stomach; Vector: puc18; Site 1: SmaI;  
Site 2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

**ORIGIN**

Query Match 1.5%; Score 46; DB 1; Length 445;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 3071 CAAGATTGTGCCACTGCTCCAGCTCGGCAACAGACGAAGACTC 3116  
|||||  
247 CAAGATTGTGCCACTGCTCCAGCTCGGCAACAGACGAAGACTC 202

**RESULT 57** AA504863 462 bp mRNA linear EST 12-AUG-1997  
AA504863  
LOCUS ab03b02.s1 StrataGene fetal retina 937202 Homo sapiens cDNA clone  
IMAGE:839691.3; similar to contains Alu repetitive element; contains  
1..53 11 repetitive element ;, mRNA sequence.

**ACCESSION** AA504863  
AA504863.1 GI:2241023

**VERSION** EST.  
AA504863.1

**KEYWORDS** Homo sapiens (human)

**SOURCE** Homo sapiens

**ORGANISM** Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.

**REFERENCE** 1 (bases 1 to 462)  
Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisler, G., Jost, S.,  
Kucaba, T., Lacy, M., Le, N., Lennon, G., Marra, M., Martin, J.,

**TITLE** Moore, B., Schellenberg, K., Steptoe, M., Tan, F., Theising, B.,  
White, Y., Wyle, T., Waterson, R. and Wilson, R.  
WashU-Merck EST Project 1997  
Unpublished (1997)

**JOURNAL** Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810

**COMMENT** Email: est@watson.wustl.edu  
This clone is available royalty-free through LNL, contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Possible reversed clone: polyA not found  
Seq primer: -40m3 fwd. ET from Amersham  
High quality sequence stop: 399.  
Location/Qualifiers  
1..462  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_1lb="IMAGE:839691"  
/sex="mixed"  
/lab\_host="SOLR (kanamycin resistant)"  
/clone\_1lb="Stratagene fetal retina 937202"  
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:  
XhoI; Cloned unidirectionally. Primer: Oligo dT. Pooled  
retinal tissue. Average insert size: 1.0 kb uni-ZAP XR  
Vector; -5' adaptor sequence: 5' GAATTCGCACAG 3' -3'  
adaptor sequence: 5' CTCGAGTTTGTCTTTTCTTTT 3'."

**ORIGIN**

Query Match 1.5%; Score 46; DB 1; Length 462;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 3071 CAAGATTGTGCCACTGCTCCAGCTCGGCAACAGACGAAGACTC 3116  
|||||  
14 CAAGATTGTGCCACTGCTCCAGCTCGGCAACAGACGAAGACTC 59

**RESULT 58** AA128941 464 bp mRNA linear EST 23-DEC-1997  
AA128941/c  
LOCUS zollc09.s1 StrataGene neuroepithelium NT2RAMI 937234 Homo sapiens  
cDNA clone IMAGE:567376.3; similar to contains Alu repetitive  
element; contains element MER22 repetitive element ;, mRNA sequence.

**ACCESSION** AA128941  
AA128941.1 GI:1688804

**VERSION** EST.  
AA128941.1

**KEYWORDS** Homo sapiens (human)

**SOURCE** Homo sapiens

**ORGANISM** Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.

**REFERENCE** 1 (bases 1 to 464)  
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapell, B.,  
Chisose, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,  
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,  
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,  
Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,  
Trevasakis, E., Underwood, K., Wohldmann, P., Waterson, R., Wilson, R.  
and Marra, M.

**TITLE** Generation and analysis of 280,000 human expressed sequence tags

**JOURNAL** Genome Res. 6 (9), 807-828 (1996)

**COMMENT** 8889549  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
WARNING: There is evidence that suggests that the 384-well parent  
plate of this clone contains both human and mouse derived clones.

Thus, the origin of this clone is uncertain. This caution should be kept in mind should you use this clone.

This clone is available royalty-free through LML; contact the IMAG Consortium (info@image.lml.gov) for further information.

Insert Length: 706 Std Error: 0.00  
Seq primer: -40M13 fwd. from Amersham

High quality sequence stop: 317.

## FEATURES

source

```
1.464
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:4591775"
/db_xref="taxon:9606"
/clone="IMAGE:567376"
/dev_stage="Ntera-2/RA+MI neuroepithelial cells"
/lab_host="SOLR (kanamycin resistant)"
/clone_1lb="Stratagene neuroepithelium NT3RAM1 937234"
/vector="Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: oligo dT. NT2 (Ntera-2/cl.D1) precursor cells induced with Retinoic Acid for 1 week, followed by 3 weeks in mitotic inhibitors (Repate #2). Average insert size: 1.1 kb; Uni-ZAP XR vector; ~5' adaptor sequence: 5' GAATTCGCGCAGAG 3' ~3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3'"
```

## ORIGIN

Query Match 1.5%; Score 46; DB 1; Length 464;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACCCTGAGCCGAGAGTTGAGACCAAGCCTGGCCAAACAT 2940  
|||||  
DB 206 GGTGATCACCCTGAGCCGAGAGTTGAGACCAAGCCTGGCCAAACAT 161

RESULT 59 482 bp DNA linear GSS 21-MAY-1999  
AQ383997  
LOCUS RPCI11-152K20.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-152K20,  
DEFINITION genomic survey sequence.  
ACCESSION AQ383997  
VERSION AQ383997.1 GI:4355020  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Zhao,S., Adams,M.D., Niemman,W., Malek,J., de Jong,P. and Venter,J.C.  
1 (bases 1 to 482)

TITLE Use of BAC End Sequences from library RPCI-11 for Sequence-Ready

JOURNAL Map Building  
COMMENT Unpublished (1997)  
Other GSSs: RPCI11-152K20.TJ  
Contact: Shaying Zhao, William Niemman, Mark Adams  
Department of Bukayotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbe@igf.org

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Piter de Jong (piter@delong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from Research Genetics (info@resgen.com). BAC end search page: http://www.igf.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html  
Seq primer: Sp6  
Class: BAC ends.  
Location/Qualifiers  
1.482

FEATURES  
source

```
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="GDB:7558243"
/db_xref="taxon:9606"
/clone="RPCI-11-152K20"
/sex="Male"
/cell_type="Lymphocytes"
/clone_1lb="RPCI-11"
/vector="Vector: pBACe3 6; Site 1: EcoRI; Site 2: EcoRI; RPCI11 Human Male BAC Library"
```

## ORIGIN

Query Match 1.5%; Score 46; DB 9; Length 482;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACCCTGAGCCGAGAGTTGAGACCAAGCCTGGCCAAACAT 2940  
|||||  
DB 420 GGTGATCACCCTGAGCCGAGAGTTGAGACCAAGCCTGGCCAAACAT 465

RESULT 60 507 bp DNA linear GSS 23-SEP-1998  
AQ132449  
LOCUS HS 3003 A2 A08 MR CIT Approved Human Genomic Sperm Library D Homo  
DEFINITION sapiens genomic clone Plate=3003 Col=16 Row=A, genomic survey  
sequence.  
ACCESSION AQ132449  
VERSION AQ132449.1 GI:3509615  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 507)  
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.  
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome  
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
10449764  
Contact: Mahairas G.G., Wallace J.C., Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu

Sequence Tagged Connector  
Plate: 3003 Row: A Column: 16  
Class: BAC ends  
High quality sequence stop: 507.  
Location/Qualifiers  
1.507

```
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="Plate=3003 Col=16 Row=A"
/sex="male"
/clone_1lb="CIT Approved Human Genomic Sperm Library D"
/vector="Vector: pBelorAC11; BAC Clones in B-Col1 DH10B"
```

## ORIGIN

Query Match 1.5%; Score 46; DB 9; Length 507;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACCCTGAGCCGAGAGTTGAGACCAAGCCTGGCCAAACAT 2940  
|||||  
DB 125 GGTGATCACCCTGAGCCGAGAGTTGAGACCAAGCCTGGCCAAACAT 170

RESULT 61  
BX641230/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BX641230 527 bp mRNA linear EST 04-SEP-2003  
DKFDP686D19134.r1.686 (synonym: h1ccc3) Homo sapiens cDNA clone  
DKFDP686D19134\_5', mRNA sequence.  
BX641230  
BX641230.1 GI:34475549  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 527)  
Poustka, A., Albert, R., Moosmayer, P., Schupp, I., Wellenreuther, R.,  
Mewes, H.W., Well, B., Amid, C., Oesanger, A., Fobo, G., Han, M. and  
Wiemann, S.  
EST (Poustka, A., Albert, R., Moosmayer, P., Schupp, I.,  
Wellenreuther, R., et al.)  
Unpublished (2003)  
Contact: MIPS  
MIPS  
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
This is the 5' sequence of the clone insert  
clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German genome project.  
No a1 sequence available.  
This clone (DKFZP686D19134) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
Location/Qualifiers  
1..527  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZP686D19134"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="686 (synonym: h1ccc3)"  
/note="vector: pTripleX2; Site\_1: SfiI; Site\_2: SfiIB;  
cDNA-collection"

ORIGIN

Query Match 1.5%; Score 46; DB 5; Length 527;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGATGATCACTGAGCCAGAGAGTTGAGACCAAGCTGCGCAACAT 2940  
|||||  
DB 502 GGTGATCACTGAGCCAGAGAGTTGAGACCAAGCTGCGCAACAT 457

RESULT 62  
BU960416  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BU960416 550 bp mRNA linear EST 21-OCT-2002  
AGENCOURT 10628034 NIH-MGC 141 Homo sapiens cDNA clone  
IMAGE:6735149 5', mRNA sequence.  
BU960416  
BU960416.1 GI:24189988  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 550)  
NIH-MGC <http://mgc.ncl.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL  
COMMENT

Unpublished (1999)  
Contact: Robert Strassberg, Ph.D.  
Email: [cgabbs-remail.nih.gov](mailto:cgabbs-remail.nih.gov)  
Tissue Procurement: NCI  
cDNA Library Preparation: Michael Brownstein Laboratory  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ILNL at:  
<http://image.llnl.gov>  
Plate: L1CM3071 row: 0 column: 20  
High quality sequence stop: 436.  
Location/Qualifiers  
1..550  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6735149"  
/tissue\_type="mixed (pool of 40 RNAs)"  
/lab\_host="DH10B (TI-phage-resistant)"  
/clone\_lib="NIH-MGC 141"  
/note="vector: pDNR-LIB; Site\_1: SfiI (ggccattatggcc);  
Site\_2: SfiI (ggccgctcgcc); Double-stranded cDNA was  
prepared from a pool of 40 cell line polyA+ RNAs (bladder  
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -  
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,  
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,  
ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary  
gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were  
used in cloning as follows:  
5'-AAGCAGTGTATCAACGACGAGAGTGGCCATTACGCGCGG-3' and  
5'-ATTCTAGAGCCGAGCGGCGGACGATG-dt(30)NN-3'. Full-length  
enriched library was constructed using the Clontech  
Creator SMART kit and size-selected to contain the 0.2-0.5  
kb size fraction (other fractions present in NIH-MGC 142).  
Library created in the laboratory of M. Brownstein (NIH,  
NIH). Note: this is a NIH-MGC Library."

ORIGIN

Query Match 1.5%; Score 46; DB 5; Length 550;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGATGATCACTGAGCCAGAGAGTTGAGACCAAGCTGCGCAACAT 2940  
|||||  
DB 297 GGTGATCACTGAGCCAGAGAGTTGAGACCAAGCTGCGCAACAT 342

RESULT 63  
CR960770/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

CR960770 564 bp DNA linear GSS 06-JUN-2005  
Homo sapiens BAC end sequence of RZPD8737F022022D from genomic  
library (orig. Pieter J. de Jong library RPCI-11), genomic survey  
sequence.  
CR960770  
CR960770.1 GI:66991336  
GSS.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 564)  
Schwarz, F., Neubert, P., Schneider, D., Peters, M. and Korn, B.  
Direct Submission  
Submitted (05-JUN-2005) RZPD Deutsches Ressourcenzentrum fuer  
Genomforschung GmbH, Im Neuenheimer Feld 515, D-69120 Heidelberg,  
Germany  
RZPD: RZPD8737F022022D; RPI1-59M4;  
derived from Pieter J. de Jong library RPCI-11;  
<http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneID=RZPD8737F022022D>  
RZPDLIB; (Human Genomic Set - RZPD 1.0) RZPD LIB No.737  
<http://www.rzpd.de/cgi-bin/products/set.cgi?libIDNo=737>

http://www.rzpd.de/products/genomicset/  
Contact: Inge Arlart  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH,  
Heubnerweg 6, D-14059 Berlin, Germany  
Tel: +49 30 32639 100  
Fax: +49 30 32639 111  
www.rzpd.de

This clone is available from RZPD;  
Contact RZPD (customer.servic@rzpd.de) for further information.  
Clone distribution: http://www.rzpd.de/products/genomicset/  
Seq-primer: T7 (TAA-TAC-GAC-TCA-CTA-TAG-GS)  
Class: BAC ends.

# FEATURES

source  
1..564  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="RZPDB737F022022D"  
/sex="Male"  
/clone\_1lb="RPCI-11"  
/note="Vector: pBACe3.6; RPCI-11 Human Male BAC Library"

# ORIGIN

Query Match 1.5%; Score 46; DB 11; Length 564;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTC 3116  
DB 221 CAAGATTGTGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTC 176

RESULT 64  
CN265832 574 bp mRNA linear EST 16-MAY-2004  
DEFINITION 17000424824545 GRN\_EB Homo sapiens cDNA 5', mRNA sequence.  
ACCESSION CN265832  
VERSION CN265832.1 GI:47282246  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
1 (bases 1 to 574)  
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Murage, J., Fisk, G.J.,  
Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M.S., Mandalam, R.,  
Lebkowski, J. and Stanton, L.W.  
Transcriptome characterization elucidates signaling networks that  
control human ES cell growth and differentiation  
Nat. Biotechnol. 22 (6), 707-716 (2004)  
1516197  
Contact: Brandenberger R  
Regenerative Medicine  
Geron Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@geron.com  
Insert Length: 574 Std Error: 0.00.  
Location/Qualifiers  
1..574  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_1lb="GRN\_EB"  
/note="Toigo dt primed, full-length enriched cDNA library  
from embryoid body outgrowth derived from hns cell lines  
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free  
conditions."

# REFERENCE

# AUTHORS

TITLE

# JOURNAL

# PUBMED

# COMMENT

# FEATURES

# source

Query Match 1.5%; Score 46; DB 7; Length 574;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTC 3116  
DB 169 CAAGATTGTGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTC 214

RESULT 65  
B95171/c 587 bp DNA linear GSS 26-JUN-1998  
DEFINITION CIT-HSP-2172H8.TF CIT-HSP Homo sapiens genomic clone 2172H8,  
genomic survey sequence.  
ACCESSION B95171  
VERSION B95171.1 GI:2977508  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
1 (bases 1 to 587)  
Adams, M.D., Rounnaley, S.D., Zhao, S., Field, C.E., Baas, S., Linher, K.,  
Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H.,  
Simon, M. and Venter, J.C.  
Use of a random BAC End Sequence Database for Sequence-Ready Map  
Building (1998)  
Unpublished (1998)  
Contact: Mark Adams  
Department of Eukaryotic Genomics  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdamas@tigr.org  
Clones are available from Research Genetics (info@reggen.com). BAC  
end search page:  
http://www.tigr.org/cdb/hungen/bac\_end\_search/bac\_end\_search.html  
Seq primer: M13-21;  
Class: BAC ends.

# REFERENCE

# AUTHORS

# TITLE

# JOURNAL

# COMMENT

The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdamas@tigr.org  
Clones are available from Research Genetics (info@reggen.com). BAC  
end search page:  
http://www.tigr.org/cdb/hungen/bac\_end\_search/bac\_end\_search.html  
Seq primer: M13-21;  
Class: BAC ends.

# FEATURES

# source

1..587  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/db\_xref="GDB:7103850"  
/clone\_1lb="CIT-HSP"  
/sex="Male"  
/cell\_type="Sperm"  
/clone\_1lb="CIT-HSP"  
/note="Vector: pBelobAC11, Site\_1: HindIII, Site\_2:  
HindIII"

# ORIGIN

Query Match 1.5%; Score 46; DB 9; Length 587;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTC 3116  
DB 564 CAAGATTGTGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTC 519

# RESULT 66

# LOCUS

AA017813 638 bp DNA linear GSS 09-JUN-1998  
DEFINITION CIT-HSP-2307L15.TR CIT-HSP Homo sapiens genomic clone 2307L15,  
genomic survey sequence.

# ACCESSION

# VERSION

AA017813 GI:3196549  
KEYWORDS GSS.

# ORIGIN

**SOURCE** Homo sapiens (human)

**ORGANISM** Homo sapiens

**REFERENCE** 1 (bases 1 to 638)  
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Baas,S., Linher,K., Golden,K., Berry,K., C.G., Zhao,S., Field,C.E., Baas,S., Linher,K., Simon,M., and Venter,J.C., Granger,D., Suh,B., Wible,C., Shizuya,H., Use of a random BAC End Sequence Database for Sequence-Ready Map Building (1998)

**TITLE** Unpublished (1998)

**JOURNAL** Other GSSs: CIT-HSP-2307L15.TF

**COMMENT** Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdamad@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC end search page:  
http://www.tigr.org/cdb/hungen/bac\_end\_search/bac\_end\_search.html.  
Seq primer: M13 Reverse  
Class: BAC ends.

**FEATURES** Location/Qualifiers  
1..638  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="2307L15"  
/sex="Male"  
/cell\_type="Sperm"  
/clone\_1lb="CIT-HSP"  
/note="Vector: pBelOBAC11; Site\_1: HindIII; Site\_2: HindIII"

**ORIGIN**

Query Match 1.5%; Score 46; DB 9; Length 638;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**Db** 3071 CAAGATTGCGCAGTCCAGCTCCAGCTGGGCAAGAGAGAGACTC 3116  
551 CAGATTGCGCAGTCCAGCTCCAGCTGGGCAAGAGAGAGACTC 506

**RESULT** 67  
BX956845 648 bp mRNA linear EST 01-MAR-2004  
LOCUS DKFZP781H1492.1 781 (synonym: hlc4) Homo sapiens cDNA clone  
DEFINITION DKFZP781H1492.1 5', mRNA sequence.  
ACCESSION BX956845  
VERSION BX956845.1 GI:43442084  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 648)  
Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R., Mewes,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.  
EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R., et al.)  
Unpublished (2003)  
Contact: MIPS  
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;

sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.  
No si sequence available.  
This clone (DKFZP781H1492) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

**FEATURES** Location/Qualifiers  
1..648  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZP781H1492"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_1lb="781 (synonym: hlc4)"  
/note="Vector: pSport1\_Sfi; Site\_1: SfiI; Site\_2: SfiI; cDNA-collection"

**ORIGIN**

Query Match 1.5%; Score 46; DB 5; Length 648;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**Db** 2890 AGCAGGTGATCAGTCCAGGCGAGGAGTTCAGACGAGCTGGCC 2935  
47 AGCAGGTGATCAGTCCAGGCGAGGAGTTCAGACGAGCTGGCC 92

**RESULT** 68  
BX641474 657 bp mRNA linear EST 04-SEP-2003  
LOCUS DKFZP686I14135.1 686 (synonym: hlc3) Homo sapiens cDNA clone  
DEFINITION DKFZP686I14135.1 5', mRNA sequence.  
ACCESSION BX641474  
VERSION BX641474.1 GI:34475793  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 657)  
Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R., Mewes,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and Wiemann,S.  
EST (Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R., et al.)  
Unpublished (2003)  
Contact: MIPS  
MIPS Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by DKFZ (German Cancer Research Center, Heidelberg/Germany) within the cDNA sequencing consortium of the German Genome Project.  
No si sequence available.  
This clone (DKFZP686I14135) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

**FEATURES** Location/Qualifiers  
1..657  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZP686I14135"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_1lb="686 (synonym: hlc3)"  
/note="Vector: pTriplex2; Site\_1: SfiI; Site\_2: SfiI; cDNA-collection"

## ORIGIN

Query Match 1.5%; Score 46; DB 5; Length 657;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATGACCTGAGGCCAGAGATTGAGACCGAGCTGGCCAAACAT 2940  
|||||  
DB 502 GGTGATGACCTGAGGCCAGAGATTGAGACCGAGCTGGCCAAACAT 457

## RESULT 69

CA430010

LOCUS 687 bp mRNA linear EST 07-NOV-2002  
DEFINITION UI-H-FH1-bfp-j-24-0-UI.61 NCI CGAP FH1 Homo sapiens cDNA clone  
CA430010  
VERSION CA430010.1 GI:24792736  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.  
1 (bases 1 to 687)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: James Martin  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained  
from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
The following repetitive elements were found in this cDNA  
sequence: 1-36, >AT\_RichLow\_complexity (matched complement)  
285-317, >AT\_RichLow\_complexity (matched complement) 369-658, >ALU  
Seq primer: M13 FORWARD  
POLYA=Yes.

## FEATURES

source

Location/Qualifiers  
1..687  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-H-FH1-bfp-j-24-0-UI"  
/tissue\_type="Cell line"  
/dev\_stage="Adult"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_1lb="NCI CGAP FH1"  
/note="Organ: Chondrosarcoma; Vector: pT73-Pac  
(pharmacia) with a modified polylinker; Site 1: EcoR I;  
Site 2: Not I; NCI CGAP FH1 is a normalized cDNA library  
obtained from a cell line derived from grade I  
chondrosarcoma tissue. The library was constructed and  
normalized according to Bonaldo, Lennon and Soares, Genome  
Research, 6:791-806, 1996. First strand cDNA synthesis was  
primed with an oligo-dT primer containing a Not I site.  
Double stranded cDNA was ligated to an EcoR I adaptor,  
digested with Not I, and cloned directionally into  
pT73-Pac vector. The oligonucleotide used to prime the  
synthesis of first-strand cDNA contains a library tag  
sequence that is located between the Not I site and the  
(dT)18 tail. The sequence tag for this library is  
ACAATCGGC. The cell line was provided by Dr. James Martin  
from the University of Iowa.  
TAG\_TISSUE=human Chondrosarcoma Cell line C8 - Grade 1  
Chondrosarcoma  
TAG\_1lb=UI-H-FH1  
TAG\_SEQ=AGAAATCCGGC"

## ORIGIN

Query Match 1.5%; Score 46; DB 6; Length 687;  
Best Local Similarity 100.0%; Pred. No. 2e-10;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATGACCTGAGGCCAGAGATTGAGACCGAGCTGGCCAAACAT 2940  
|||||  
DB 424 GGTGATGACCTGAGGCCAGAGATTGAGACCGAGCTGGCCAAACAT 469

## RESULT 70

BUSE64032

LOCUS 258 bp mRNA linear EST 16-SEP-2002  
DEFINITION AGENCOURT 10371369 NIH MGC 141 Homo sapiens cDNA clone  
IMAGE:6601889 5', mRNA sequence.  
ACCESSION BUSE64032  
VERSION BUSE64032.1 GI:22914332  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominiidae; Homo.  
1 (bases 1 to 258)  
NIH-MGC <http://mhc.nci.nih.gov/>.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: NCI  
CDNA Library Preparation: Michael Brownstein Laboratory  
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)  
Clone Distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
plate: LNCM2840 row: d column: 17  
High quality sequence stop: 171.

## FEATURES

source

Location/Qualifiers  
1..258  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6601889"  
/tissue\_type="mixed (pool of 40 RNAs)"  
/lab\_host="DH10B (Tl-phage-resistant)"  
/clone\_1lb="NIH MGC 141"  
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggc);  
Site 2: SfiI (ggccgctcgcc); Double-stranded cDNA was  
prepared from a pool of 40 cell line polyA+ RNAs (bladder  
2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -  
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,  
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,  
ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary  
gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were  
used in cloning as follows:  
5'-AAGCAGGTATGACGAGCGGATGCACTTACGGCCGG-3' and  
5'-ATTGAGAGCCGAGCGGCGACATG-dt(30)NN-3'. Full-length  
enriched library was constructed using the Clontech  
Creator SMART kit and size-selected to contain the 0.2-0.5  
kb size fraction (other fractions present in NIH MGC 142).  
Library created in the laboratory of M. Brownstein (NIH),  
NIH. Note: this is a NIH\_MGC library."

## ORIGIN

Query Match 1.4%; Score 45; DB 5; Length 258;  
Best Local Similarity 100.0%; Pred. No. 6.2e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGGCACTGCACTCGAGGCTGGGCAAGAGCAAGACT 3115  
|||||  
DB 174 CAAGATTGGCACTGCACTCGAGGCTGGGCAAGAGCAAGACT 218

RESULT 71  
LOCUS BX953709 287 bp mRNA linear EST 01-MAR-2004  
DEFINITION DKFZP781C034.r1.781 (synonym: h1cc4) Homo sapiens cDNA clone  
DKFZP781C034.5', mRNA sequence.  
ACCESSION BX953709  
VERSION BX953709.1 GI:43433046  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 287)  
AUTHORS Ansoerge,W., Krieger,S., Regiert,T., Rittmuller,C., Schwager,B.,  
Mewes,H.W., Weill,B., Amid,C., Oesanger,A., Fobo,G., Han,M. and  
Wiemann,S.  
TITLE EST (Ansoerge W., Krieger S., Regiert T., Rittmuller C., et al.)  
JOURNAL Unpublished (2003)  
COMMENT Contact: MIPS  
MIPS  
Ingolstaedter Landstr.1, D-85764 Neuberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ), Email s.wiemann@dkfz-heidelberg.de;  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
No sl sequence available.  
This clone (DKFZP781C034) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubneweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
Location/Qualifiers  
source  
1..287  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZP781C034"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_1ib="781 (synonym: h1cc4)"  
/note="Vector: pSPORT1\_Sf1; Site\_1: Sf1A; Site\_2: Sf1B;  
cDNA-collection"

ORIGIN  
Query Match 1.4%; Score 45; DB 5; Length 287;  
Best Local Similarity 100.0%; Pred. No. 6.2e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2898 GGATCCTGAGCCGAGGTTCCAGACCAAGCTGGCCAACTAG 2942  
|||||  
DB 241 GGATCCTGAGCCGAGGTTCCAGACCAAGCTGGCCAACTAG 285

RESULT 72  
LOCUS AQ485964 309 bp DNA linear GSS 24-APR-1999  
DEFINITION RPCI-11-23519.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-23519,  
genomic survey sequence.  
ACCESSION AQ485964  
VERSION AQ485964.1 GI:4668012  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 309)  
AUTHORS Zhao,S., Adams,M.D., Niernan,W., Malek,J., de Jong,P. and  
Venter,J.C.  
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
Map Building

JOURNAL Unpublished (1997)  
COMMENT Contact: Shaying Zhao, William Niernan, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbe@tigr.org  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieter@dejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
Research Genet cs (info@resgen.com). BAC end search page:  
http://www.tigr.org/tcdb/humgen/bac\_end\_search/bac\_end\_search.html.  
Seq primer: Sp6  
Class: BAC ends.  
Location/Qualifiers  
source  
1..309  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7590056"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-23519"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_1ib="RPCI-11"  
/note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
RPCI11 Human Male BAC Library"

ORIGIN  
Query Match 1.4%; Score 45; DB 9; Length 309;  
Best Local Similarity 100.0%; Pred. No. 6.1e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2898 GGATCCTGAGCCGAGGTTCCAGACCAAGCTGGCCAACTAG 2942  
|||||  
DB 209 GGATCCTGAGCCGAGGTTCCAGACCAAGCTGGCCAACTAG 165

RESULT 73  
LOCUS AQ381523 352 bp DNA linear GSS 21-MAY-1999  
DEFINITION RPCI11-164A21.TV RPCI-11 Homo sapiens genomic clone RPCI-11-164A21,  
genomic survey sequence.  
ACCESSION AQ381523  
VERSION AQ381523.1 GI:4352546  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 352)  
AUTHORS Zhao,S., Adams,M.D., Niernan,W., Malek,J., de Jong,P. and  
Venter,J.C.  
TITLE Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
Map Building (1997)  
JOURNAL Unpublished (1997)  
COMMENT Other GSSs: RPCI11-164A21.TJ  
Contact: Shaying Zhao, William Niernan, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbe@tigr.org  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieter@dejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
Research Genetcs (info@resgen.com). BAC end search page:  
http://www.tigr.org/tcdb/humgen/bac\_end\_search/bac\_end\_search.html  
Seq primer: T7



Class: BAC ends.  
Location/Qualifiers  
1.352  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7562612"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-164A21"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_11b="RPCI-11"  
/note="Vector: pBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
RPC11 Human Male BAC Library"

Query Match 1.4%; Score 45; DB 9; Length 352;  
Best Local Similarity 100.0%; Pred. No. 6.1e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCACTGCACTCCAGCTGGGCAACAGAGCAAGACT 3115  
DB 145 CAAGATTGTGCACTGCACTCCAGCTGGGCAACAGAGCAAGACT 189

RESULT 74  
LOCUS AQ485959 408 bp DNA linear GSS 24-APR-1999  
DEFINITION RPCI-11-23515.TV RPCI-11 Homo sapiens genomic clone RPCI-11-23515,  
genomic survey sequence.  
ACCESSION AQ485959  
VERSION AQ485959.1 GI:4668007  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 408)  
Zhao,S., Adams,M.D., Niernan,W., Malek,J., de Jong,P. and  
Vener,J.C.  
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
Map Building  
Unpublished (1997)  
Other GSSs: RPCI-11-23515.TV  
Contact: Shaying Zhao, William Niernan, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbeet@igir.org  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pdejong@med.buhalo.edu). Clones may be purchased from  
BACAC Resources (http://bacpac.med.buhalo.edu/ordering) or from  
Research Genet cs (http://resgen.com). BAC end search page:  
http://www.igir.org/igir/humgen/bac\_end\_search/bac\_end\_search.html.  
Seq primer: Sp6  
Class: BAC ends.  
Location/Qualifiers  
1.408  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7590052"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-23515"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_11b="RPCI-11"  
/note="Vector: pBACE3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
RPC11 Human Male BAC Library"

ORIGIN  
source

Query Match 1.4%; Score 45; DB 9; Length 408;  
Best Local Similarity 100.0%; Pred. No. 6.1e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2898 GATCATCTGAGCCAGAGATTGAGACCAAGCTTGGCCCAATAG 2942  
DB 306 GATCATCTGAGCCAGAGATTGAGACCAAGCTTGGCCCAATAG 262

RESULT 75  
LOCUS BU951869 415 bp mRNA linear EST 21-OCT-2002  
DEFINITION BU951869.x1 H85 Islet Homo sapiens cDNA clone IMAGE:6127759 3',  
mRNA sequence.  
ACCESSION BU951869.1 GI:24203621  
VERSION BU951869.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 415)  
Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,  
Lemishka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,  
Hillier,L., Maira,M., Page,D., Wylie,T., Martin,J., Blaisstein,A.,  
Schmitt,A., Thelting,B., Ritter,B., Ronko,I., Bennett,J.,  
Cardenas,M., Gibbons,M., McCann,R., Cole,R., Teagareishvili,R.,  
Williams,T., Jackson,Y. and Bowers,Y.  
Endocrine Pancreas Consortium  
Unpublished (2000)  
Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
Endocrine Pancreas Consortium  
Harvard University, Howard Hughes Medical Institute  
Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
MA 02138  
Tel: 617-495-1812  
Fax: 617-495-8557  
Email: dmelton@bcb.harvard.edu  
Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
Washington University Genome Sequencing Center For information on  
obtaining a clone please contact: Dr. Hiroshi Inoue  
(hinoue@im.wustl.edu)  
Seq primer: -40up from Gibco  
High quality sequence stop: 337.  
Location/Qualifiers  
1.415  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6127759"  
/tissue\_type="Purified pancreatic islet"  
/lab\_host="DH10B"  
/clone\_11b="H85 Islet"  
/note="Organ: Pancreas; Vector: pBluescript SK(-); Site\_1:  
NotI; Site\_2: XhoI; cDNA made by oligo-dT priming.  
Size: selected on agarose gel. Average insert size -1kb. 5'  
XhoI site was destroyed after directional cloning.  
Amplified once. Contact information: Hiroshi Inoue, MD,  
Metabolism Div. (Alan Permutt Lab), Washington University  
School of Medicine, Box 8127, 660 South Euclid Ave., St.  
Louis, MO 63110, E-mail: hinoue@imgate.wustl.edu, Tel:  
314-362-1916, Fax: 314-747-2692."

ORIGIN  
source

Query Match 1.4%; Score 45; DB 5; Length 415;  
Best Local Similarity 100.0%; Pred. No. 6.1e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCACTGCACTCCAGCTGGGCAACAGAGCAAGACT 3115  
DB 51 CAAGATTGTGCACTGCACTCCAGCTGGGCAACAGAGCAAGACT 7

RESULT 76  
LOCUS A0065412/c 417 bp DNA linear GSS 04-AUG-1998  
DEFINITION HS 2224 B1 P05 MR CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate=2224 Col=9 Row=L, genomic survey sequence.  
ACCESSION A0065412  
VERSION A0065412.1 GI:3380930  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 417)  
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.  
High throughput sequencing  
Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Sequence Tagged Connector  
Plate: 2224 row: L column: 9  
Class: BAC ends  
High quality sequence stop: 417.  
Location/Qualifiers  
1..417  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="Plate=2224 Col=9 Row=L"  
/sex="male"  
/clone\_1ib="CIT Approved Human Genomic Sperm Library D"  
/note="Organ: sperm; Vector: pBelobAC11; BAC Clones in E-Coli DH10B"

ORIGIN  
Query Match 1.4%; Score 45; DB 9; Length 417;  
Best Local Similarity 100.0%; Pred. No. 6.1e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
3078 GTGGCACTGCACTCCAGCTGGGCAACAGCAAGACTCTGTCTC 3122  
|||||  
414 GTGGCACTGCACTCCAGCTGGGCAACAGCAAGACTCTGTCTC 370

RESULT 77  
LOCUS AA085683/c 424 bp mRNA linear EST 01-DEC-1996  
DEFINITION zn53d01.s1 StrataGene musclic 937209 Homo sapiens cDNA clone IMAGE:561889 3' similar to contains Alu repetitive element; , mRNA sequence.  
ACCESSION AA085683  
VERSION AA085683.1 GI:1629145  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 424)  
Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chappellet,B., Chisoe,S., Dietrich,N., Dubuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N.,

MAHAIAS,B., MOORE,B., MORRIS,M., PARSONS,J., PRANGE,C., RIFKIN,L., ROHLFING,T., SCHALLENBERG,K., SOARES,M.B., TAN,F., THIERRY-MEG,J., TREVASKIS,B., UNDERWOOD,K., WOHLHAMM,P., WATERSON,R., WILSON,R. and MATRIZ,M.  
Generation and analysis of 280,000 human expressed sequence tags  
Genome Res. 6 (9), 807-828 (1996)  
889549  
Contact: Wilson RK  
Washington University School of Medicine  
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Insert Length: 2224 Std Error: 0.00  
Seq primer: -40M13 fwd. from Amersham  
High quality sequence stop: 393.  
Location/Qualifiers  
1..424  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:4595179"  
/db\_xref="taxon:9606"  
/clone="IMAGE:561889"  
/tissue\_type="muscle"  
/dev\_stage="adult"  
/lab\_host="SOLR (kanamycin resistant)"  
/clone\_1ib="Stratagene muscle 937209"  
/note="Organ: skeletal muscle; Vector: pBluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Skeletal muscle from patient with malignant hyperthermia. Average insert size: 1.0 kb; Uni-ZAP XR Vector; -5' adaptor sequence: 5' GAATTCGACGACGAG 3' -3' adaptor sequence: 5' CTCGAGTTTCTTTTCTTTT 3' "

ORIGIN  
Query Match 1.4%; Score 45; DB 1; Length 424;  
Best Local Similarity 100.0%; Pred. No. 6.1e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
2888 TGAGCAGGTGGATCCTGAGCCAGAGGTTGAGACAGGCTG 2932  
|||||  
222 TGAGCAGGTGGATCCTGAGCCAGAGGTTGAGACAGGCTG 178

RESULT 78  
LOCUS CB132267/c 436 bp mRNA linear EST 29-JAN-2003  
DEFINITION K-EST0182497 L11SN035481 Homo sapiens cDNA clone L11SN035481-1-B08 5' mRNA sequence.  
ACCESSION CB132267  
VERSION CB132267.1 GI:28097454  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 436)  
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R., Oh,K.U., Cheong,J.B., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and Kim,Y.S.  
21C Frontier Korean EST Project 2001  
Unpublished (2002)  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-383, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 1 row: E column: 08

TITLE  
JOURNAL  
COMMENT

FEATURES  
source  
High quality sequence stop: 436.  
Location/Qualifiers  
1.436  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="U11SNJ5481-1-E08"  
/sex="M"  
/tissue\_type="Liver"  
/cell\_type="Polygonal"  
/cell\_line="SNJ-354"  
/lab\_host="Top10F"  
/clone\_lib="U11SNJ5481"  
/note="Organ: Liver; Vector: pCNS-D2; Site 1: EcoRI; Site 2: NotI; The poly (A)+ RNA was dephosphorylated with bacterial alkaline phosphatase (BAP) and then deprotected with tobacco acid pyrophosphatase (TAP). The deprotected intact mRNA was ligated with DNA-RNA linker including EcoRI site by treatment of T4 RNA ligase and the first strand cDNA was synthesized from oligo dT-selected mRNA by priming with dT-tailed vector. The dT-tailed vector was adjusted to have about 60nt. The cDNA vector was circularized with E. coli DNA ligase after digestion of EcoRI which site is also included in vector. An RNA strand converted to a DNA strand by Okayama-Berg method. The obtained cDNA vectors were used for transformation of competent cells E. coli Top10F by electroporation method. The cDNA libraries constructed by this method are full-length enriched cDNA library. After analyzing and sequencing about 2,000 - 3,000 colonies in original cDNA library, the abundant cDNAs were selected and amplified by PCR reaction using vector region primer including T7 promoter as 5' primer and NID14 as 3' primer. The PCR products were used as template for synthesis of biotinylated single stranded RNA by in vitro transcription reaction. The synthesized RNA probes were hybridized with antisense single stranded cDNAs prepared from original library and incubated with avidin-gel. After removing DNA-RNA hybrids by centrifuge, the substracted cDNA libraries were constructed by transformation of the remaining DNA into competent cells E. coli Top10F with electroporation method."

ORIGIN  
Query Match 1.4%; Score 45; DB 6; Length 436;  
Best Local Similarity 100.0%; Pred. No. 6.1e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2895 GGTGATCAGCTGAGCGCAGAGTTGAGACCGCTGGCCACA 2939  
|||||  
233 GGTGATCAGCTGAGCGCAGAGTTGAGACCGCTGGCCACA 189

RESULT 79  
LOCUS A0009550 439 bp DNA linear GSS 27-JUN-1998  
DEFINITION CIT-HSP-2285K1.TRB CIT-HSP Homo sapiens genomic clone 2285K1,  
genomic survey sequence.  
ACCESSION A0009550  
VERSION A0009550.1 GI:3128935  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 439)  
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.B., Baas,S., Linher,K.,  
Simon,M., and Venter,U.C.  
Use of a random BAC End Sequence Database for Sequence-Ready Map  
Building (1998)  
JOURNAL Unpublished (1998)

COMMENT  
Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdamad@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC  
end search page:  
http://www.tigr.org/cdb/hungen/bac\_end\_search/bac\_end\_search.html  
Seq primer: M13 Reverse  
Classes: BAC ends.

FEATURES  
source  
Location/Qualifiers  
1.439  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GBD:714812"  
/db\_xref="taxon:9606"  
/clone="2285K1"  
/sex="Male"  
/cell\_type="Sperm"  
/clone\_lib="CIT-HSP"  
/note="Vector: pBel0BAC11; Site\_1: HindIII; Site\_2:  
HindIII"

ORIGIN  
Query Match 1.4%; Score 45; DB 9; Length 439;  
Best Local Similarity 100.0%; Pred. No. 6.1e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3073 AAGTTGGCACTGACCTGACCTGGGCGACAGAGCAAGACTCT 3117  
|||||  
282 AAGTTGGCACTGACCTGACCTGGGCGACAGAGCAAGACTCT 326

RESULT 80  
LOCUS A0554616/c 447 bp DNA linear GSS 28-MAY-1999  
DEFINITION RPCI-11-409A7.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-409A7,  
genomic survey sequence.  
ACCESSION A0554616  
VERSION A0554616.1 GI:4913793  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 447)  
Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and  
Venter,U.C.  
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
Map Building  
Unpublished (1997)  
Other GSSes: RPCI-11-409A7.TV  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbeet@tigr.org  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieter@delong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
Research Genet cs (info@resgen.com). BAC end search page:  
http://www.tigr.org/cdb/hungen/bac\_end\_search/bac\_end\_search.html.  
Seq primer: 586  
Classes: BAC ends.  
Location/Qualifiers  
1.447  
/organism="Homo sapiens"

FEATURES  
source

/mol\_type="genomic DNA"  
/db\_xref="GDB:765678"  
/db\_xref="taxon:9606"  
/clone="RPCT-11-409A7"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_1lb="RPCT-11"  
/note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
RPCT11 Human Male BAC Library"

## ORIGIN

Query Match 1.4%; Score 45; DB 9; Length 447;  
Best Local Similarity 100.0%; Pred. No. 6.1e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGGCAGTGCAGCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
|||||  
DB 182 GTGGCAGTGCAGCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 138

## RESULT 81

LOCUS AQ170811 452 bp DNA linear GSS 16-OCT-1998  
DEFINITION HS 3071 A1 H09 MR CIT Approved Human Genomic Sperm Library D Homo  
sapiens genomic clone plate=3071 Col=17 Row=O, genomic survey  
sequence.

ACCESSION AQ170811 GI:3568178  
VERSION AQ170811.1  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 452)  
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,  
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and  
Hood,L.

## REFERENCE

AUTHORS

TITLE Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
PUBMED 10449764  
COMMENT Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Sequence Tagged Connector  
Plate: 3071 Row: O Column: 17  
Class: BAC ends  
High quality sequence stop: 452.

## FEATURES

source

1..452  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="Plate=3071 Col=17 Row=O"  
/sex="male"  
/clone\_1lb="CIT Approved Human Genomic Sperm Library D"  
/note="Organ: sperm; Vector: pBACe3.6; BAC Clones in  
E-Coli DH10B"

## ORIGIN

Query Match 1.4%; Score 45; DB 9; Length 452;  
Best Local Similarity 100.0%; Pred. No. 6.1e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2896 GTGGATCAGCTGAGCCAGAGGTTGAGACCAAGCTGGCCAAAT 2940  
|||||  
DB 201 GTGGATCAGCTGAGCCAGAGGTTGAGACCAAGCTGGCCAAAT 245

RESULT 82  
LOCUS BG944968 469 bp mRNA linear EST 15-JAN-2003  
DEFINITION ax57a12.x1 Hembase; Erythroid Progenitor Cells (LGB:ax 1library)  
Homo sapiens CDNA clone ax57a12 random, mRNA sequence.  
BG944968  
ACCESSION BG944968.1 GI:14344340  
VERSION EST.  
KEYWORDS  
SOURCE Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 469)  
Gubin,A.N., Njoroge,J.N., Bouffard,G.G. and Miller,J.L.  
Gene expression in proliferating human erythroid cells  
Genomics 59 (2), 168-177 (1999)  
10409428

## COMMENT

Contact: Jeffery L. Miller  
Laboratory of Chemical Biology  
National Institute of Diabetes and Digestive and Kidney Diseases  
Building 10, Room 9B17, National Institutes of Health, Bethesda, MD  
20892, USA  
Tel: 301 402 2373  
Fax: 301 435 5148  
Email: jmf@nhi.gov  
DNA Sequencing and analyses by National Institutes of Health  
Intramural Sequencing Center (NISC).  
Plate: 57 Row: A Column: 12  
Seq primer: -21M13 forward primer (ABI).  
Location/Qualifiers

## FEATURES

source

1..469  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="ax57a12"  
/sex="unknown"  
/tissue\_type="blood"  
/cell\_type="Erythroid Cells"  
/cell\_line="Primary Culture of Peripheral Blood  
Mononuclear Cells"  
/dev stage="Progenitor; EPO responsive CD71++++"  
/lab host="SOLR"  
/clone\_1lb="Hembase; Erythroid Progenitor Cells (LGB:ax  
1library)"  
/note="Organ: blood; Vector: Lambda ZAP II; Site\_1: EcoRI;  
Site\_2: EcoRI; 65,000 proliferating erythroid cells from  
the buffy coat of a blood donation were obtained by flow  
cytometric separation after a 5-day culture period in the  
presence of erythropoietin. Total RNA was purified from  
the sorted cell population using TRIzol reagent. RNA (0.3  
ug) was converted into double stranded cDNA using  
Clontech's Capfinder cDNA library Construction Kit  
(Clontech) according to the manufacturer's protocol and  
cloned into EcoRI digested lambda Zap II vector  
(Stratagene). The phage library was amplified once prior  
to in vivo excision in SOLR cells. Individual colonies  
were grown, and the cDNA inserts were sequenced in high  
throughput (NISC Intramural sequencing center  
<http://www.nisc.nih.gov/>)."

## ORIGIN

Query Match 1.4%; Score 45; DB 2; Length 469;  
Best Local Similarity 100.0%; Pred. No. 6.1e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGGCAGTGCAGCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
|||||  
DB 132 GTGGCAGTGCAGCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 176

## RESULT 83

H63465/c  
LOCUS H63465 472 bp mRNA linear EST 11-OCT-1995  
DEFINITION YR53606.r1 Soares fetal liver spleen INPIS Homo sapiens cDNA clone  
IMAGE:209026 5' similar to contains Alu repetitive element; mRNA  
sequence.  
ACCESSION H63465  
VERSION H63465.1 GI:1018266  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 472)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkes, M.,  
Holman, M., Hultman, M., Kucada, T., Le, M., Lennon, G., Maitra, M.,  
Parsons, J., Rikkin, L., Rohlfing, T., Soares, M., Tan, F.,  
Trevisakis, E., Waterston, R., Williamson, A., Wohlmann, P. and  
Wilson, R.  
The MashU-Merck EST Project  
TITLE The MashU-Merck EST Project  
JOURNAL Unpublished (1995)  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@watson.wustl.edu  
Insert Size: 1442  
High quality sequence stops: 426  
Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.lnl.gov) for further information.  
Insert Length: 1442 Std Error: 0.00  
Seq primer: M13RP1  
High quality sequence stop: 426.  
Location/Qualifiers  
1..472  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/db\_xref="IMAGE:209026"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares fetal liver spleen INPIS"  
/note="Organ: Liver and Spleen; Vector: pT73D (pharmacia)  
with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;  
1st strand cDNA was primed with a Pac I - oligo(dT) primer  
[5' AACGAGAGATTAATTAAGATCTTTTCTTTTCTTTTCTT 3'],  
double-stranded cDNA was ligated to Eco RI adaptors  
(pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified pT73 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M. Patricia Bonaldo."

ORIGIN  
Query Match 1.4%; Score 45; DB 8; Length 472;  
Best Local Similarity 100.0%; Pred. No. 6.1e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3078 GTGGCAGCTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
DB 282 GTGGCAGCTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 238

RESULT 84  
LOCUS BX954324 494 bp mRNA linear EST 01-MAR-2004  
DEFINITION DKFZp781B2254\_r1 781 (synonym: h1cc4) Homo sapiens cDNA clone  
BX954324 mRNA sequence.  
ACCESSION BX954324  
VERSION BX954324.1 GI:43434818

KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 494)  
Koehrer, K., Beyer, A., Mewes, H.W., Well, B., Amlid, C., Osanger, A.,  
Rebo, G., Han, M. and Wiemann, S.  
EST (Koehrer, K., Beyer, A., Mewes, H.W., Well, B., Amlid, C., et al.)  
Unpublished (2003)  
CONTACT: MIPS  
MIPS  
Ingolstaedter landstr.1, D-85764 Neuberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by BMFZ (Biomedical Research Center at the Heinrich-  
Heine-University, Dueseldorf/Germany) within the cDNA sequencing  
consortium of the German Genome Project. No sl sequence available.  
This clone (DKFZp781B2254) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
Location/Qualifiers  
1..494  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp781B2254"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="781 (synonym: h1cc4)"  
/note="Vector: pSport1\_Sfi; Site\_1: SfiIA; Site\_2: SfiIB;  
cDNA-collection"

ORIGIN  
Query Match 1.4%; Score 45; DB 5; Length 494;  
Best Local Similarity 100.0%; Pred. No. 6.1e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3078 GTGGCAGCTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
DB 56 GTGGCAGCTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 100

RESULT 85  
LOCUS AQ038737/c 503 bp DNA linear GSS 11-JUL-1998  
DEFINITION CIT-HSP-232509.TV CIT-HSP Homo sapiens genomic clone 232509,  
genomic survey sequence.  
ACCESSION AQ038737  
VERSION AQ038737.1 GI:3304569  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominoidea; Homo.  
REFERENCE 1 (bases 1 to 503)  
Adams, M.D., Reunaley, S.D., Zhao, S., Field, C.E., Baas, S., Linher, K.,  
Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H.,  
Simon, M. and Venter, J.C.  
Use of a random BAC End Sequence Database for Sequence-Ready Map  
Building (1998)  
Unpublished (1998)  
CONTACT: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdadams@tigr.org

Clones are available from Research Genetics (info@resgen.com). BAC end search page: [http://www.cligr.org/cdb/hungen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.cligr.org/cdb/hungen/bac_end_search/bac_end_search.html). Seq primer: T7  
Class: BAC ends.

# FEATURES

source  
1..503  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="232509"  
/sex="Male"  
/cell\_type="Sperm"  
/clone\_1ib="CIT-HSP"  
/note="Vector: pBeloBAC11; Site\_1: HindIII; Site\_2: HindIII"

# ORIGIN

Query Match 1.4%; Score 45; DB 9; Length 503;  
Best Local Similarity 100.0%; Pred.No. 6.1e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTGTCTC 3122  
|||||  
Db 268 GTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTGTCTC 224

RESULT 86 CBI42858 529 bp mRNA linear EST 29-JAN-2003  
LOCUS K-EST0196755 L11SNUS481 Homo sapiens cDNA clone L11SNUS481-22-H11  
DEFINITION 5' mRNA sequence.  
ACCESSION CBI42858  
VERSION CBI42858  
KEYWORDS EST.  
SOURCE CBI42858.1 GI:28119424  
ORGANISM Homo sapiens (human)  
Homosapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 529)  
Kim,N.S., Hahn,Y., Oh,J.H., Lee,J.Y., Ahn,H.Y., Chu,M.Y., Kim,M.R.,  
Oh,K.J., Cheong,J.E., Sohn,H.Y., Kim,J.M., Park,H.S., Kim,S. and  
Kim,Y.S.  
21C Frontier Korean EST Project 2001  
Unpublished (2002)  
Contact: Kim YS  
Genome Research Center  
Korea Research Institute of Bioscience & Biotechnology  
52 Boeun-dong Yuseong-gu, Daejeon 305-333, South Korea  
Tel: +82-42-860-4470  
Fax: +82-42-860-4409  
Email: yongsung@mail.kribb.re.kr  
Plate: 22 row: H column: 11  
High quality sequence stop: 529.

# FEATURES

source  
1..529  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="L11SNUS481-22-H11"  
/sex="M"  
/issue\_type="Liver"  
/cell\_type="Polygonal"  
/cell\_line="SNU-354"  
/lab\_host="Top10F"  
/clone\_1ib="L11SNUS481"  
/note="Organ: Liver; Vector: PCNS-D2; Site\_1: EcoRI;  
Site\_2: NotI; The poly (A)+ RNA was dephosphorylated with  
bacterial alkaline phosphatase (BAP) and then decapped  
with tobacco acid pyrophosphatase (TAP). The decapped  
intact mRNA was ligated with DNA-RNA linker including  
EcoRI site by treatment of T4 RNA ligase and the first

strand cDNA was synthesized from oligo dT-selected mRNA by  
priming with dT-tailed vector. The dT-tailed vector was  
adjusted to have about 60nt. The cDNA vector was  
circularized with E. coli DNA ligase after digestion of  
EcoRI which site is also included in vector. An RNA strand  
converted to a DNA strand by Okayama-Berg method. The  
obtained cDNA vectors were used for transformation of  
competent cells E. coli Top10F by electroporation method.  
The cDNA libraries constructed by this method are  
full-length enriched cDNA library. After analyzing and  
sequencing about 2,000 - 3,000 colonies in original cDNA  
library, the abundant cDNAs were selected and amplified by  
PCR reaction using vector region primer including T7  
promotor as 5' primer and Nidrl14 as 3' primer. The PCR  
products were used as template for synthesis of  
biocynylated single stranded RNA by in vitro transcription  
reaction. The synthesized RNA probes were hybridized with  
antisense single stranded cDNAs prepared from original  
library and incubated with avidin-gel. After removing  
DNA-RNA hybrids by centrifuge, the subtracted cDNA  
libraries were constructed by transformation of the  
remaining DNA into competent cells E. coli Top10F with  
electroporation method."

# ORIGIN

Query Match 1.4%; Score 45; DB 6; Length 529;  
Best Local Similarity 100.0%; Pred.No. 6.1e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2895 GGTGATGACCTGAGCGCCAGAGTTGAGACCAAGCTGGCCACA 2339  
|||||  
Db 223 GGTGATGACCTGAGCGCCAGAGTTGAGACCAAGCTGGCCACA 189

RESULT 87 AQ333702 529 bp DNA linear GSS 06-MAR-1999  
LOCUS AQ333702/c  
DEFINITION HS 5013 A2 F10 T7 RPII1 Human Male BAC Library Homo sapiens  
genomic clone Plate=589 Col=20 Row=K, genomic survey sequence.  
ACCESSION AQ333702  
VERSION AQ333702.1 GI:4131189  
KEYWORDS GSS.  
SOURCE AQ333702.1 GI:4131189  
ORGANISM Homo sapiens (human)  
Homosapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 529)  
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,  
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and  
Hood,L.  
Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
10449764  
Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones may be purchased from Research Genetics (info@resgen.com).  
BAC end Web Server: <http://www.htsc.washington.edu>  
Plate: 589 row: K column: 20  
Seq primer: T7  
Class: BAC ends  
High quality sequence stop: 529.

TITLE  
JOURNAL  
PUBMED  
COMMENT

# FEATURES

source  
1..529  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"





FEATURES  
source  
1. 627  
Location/Qualifiers  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone="PTB-062C22.R"  
/sex="male"  
/cell\_type="Lymphoblast"  
/clone\_1ib="PTB Chimpanzee Male BAC library"

ORIGIN  
Query Match 1.4%; Score 45; DB 10; Length 627;  
Best Local Similarity 100.0%; Pred. No. 6e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 GTGCACCTGCACTCCAGCCTGGGCAACAGACGAACTCTGTCTC 3122  
|||||  
Db 225 GTGCACCTGCACTCCAGCCTGGGCAACAGACGAACTCTGTCTC 269  
|||||

RESULT 91  
AG116078/c 640 bp DNA linear GSS 03-NOV-2001  
LOCUS Pan troglodytes DNA, clone: PTB-123103.R, genomic survey sequence.  
DEFINITION AG116078  
ACCESSION AG116078.1 GI:16736597  
VERSION  
KEYWORDS  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Pan.  
REFERENCE  
1 Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,  
Toto, Y., Watanabe, H. and Sakaki, Y.  
BAC end sequences of library PTB  
Unpublished  
2 (bases 1 to 640)  
Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T.,  
Toto, Y., Watanabe, H. and Sakaki, Y.  
Direct Submission  
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC),  
1-7-22 Suenho-chou Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail:chimbasesgsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/  
Tel:81-45-503-9111, Fax:81-45-503-9170)  
Clones are derived from the chimpanzee BAC library PTB this BAC end  
was generated during the Rad process and may have higher chance of  
clone tracking errors.  
PRIMERS  
Sequencing: M13Rev  
LIBRARY  
Vector : pKS145  
R.site 1 : SacI  
R.site 2 : SacI  
Location/Qualifiers  
1. 640  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone="PTB-123103.R"  
/sex="male"  
/cell\_type="Lymphoblast"  
/clone\_1ib="PTB Chimpanzee Male BAC library"

ORIGIN  
Query Match 1.4%; Score 45; DB 10; Length 640;  
Best Local Similarity 100.0%; Pred. No. 6e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCACTGCACTCCAGCCTGGGCAACAGACGAACT 3115  
|||||  
Db 386 CAAGATTGTGCACTGCACTCCAGCCTGGGCAACAGACGAACT 342  
|||||

RESULT 92  
A0540193/c 641 bp DNA linear GSS 19-MAY-1999  
LOCUS RPCI-11-345P21.TV RPCI-11 Homo sapiens genomic clone  
DEFINITION RPCI-11-345P21, genomic survey sequence.  
ACCESSION A0540193  
VERSION A0540193.1 GI:4870723  
KEYWORDS  
SOURCE Homo sapiens (human)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE  
1 (bases 1 to 641)  
Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and  
Venter, J.C.  
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
Map Building  
Unpublished (1997)  
Other\_GSSs: RPCI-11-345P21.TV  
Contact: Shaying Zhao, William Nierman, Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbs@tigr.org  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Peter de Jong  
(peterdejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering) or from  
Research Genet cs (info@resgen.com). BAC end search page:  
http://www.tigr.org/cdb/humgen/bac\_end\_search/bac\_end\_search.html.  
Seq primer: SP6  
Class: BAC ends.  
Location/Qualifiers  
1. 641  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7632476"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-345P21"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_1ib="RPCI-11"  
/note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI;  
RPCI11 Human Male BAC library"

ORIGIN  
Query Match 1.4%; Score 45; DB 9; Length 641;  
Best Local Similarity 100.0%; Pred. No. 6e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 GTGCACCTGCACTCCAGCCTGGGCAACAGACGAACTCTGTCTC 3122  
|||||  
Db 492 GTGCACCTGCACTCCAGCCTGGGCAACAGACGAACTCTGTCTC 448  
|||||

RESULT 93  
AG174909/c 680 bp DNA linear GSS 09-JAN-2002  
LOCUS Pan troglodytes DNA, clone: R43-045K01.T7, genomic survey  
sequence.  
ACCESSION AG174909  
VERSION AG174909.1 GI:16704589  
KEYWORDS  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Pan.

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="embryonic stem cells, DMSO-treated H9 cell line"  
 /clone\_1fb="GRN PREHEP"  
 /note="Toigo dT primed, full-length enriched cDNA library from DMSO-treated hES cell line H9 (p22) maintained in feeder-free conditions"

Query Match 1.4% Score 45; DB 7; Length 712;  
 Best Local Similarity 100.0%; Pred. No. 66-10;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2896 GTGATCACCCTGAGCCAGGAGTTTCGAGACACCTGCGCCACAT 2940  
 |||||  
 319 GTGATCACCCTGAGCCAGGAGTTTCGAGACACCTGCGCCACAT 363

RESULT 95  
 A0349458  
 LOCUS RPCI11-118J16.TJ RPCI-11 Homo sapiens genomic clone RPCI-11-118J16,  
 DEFINITION genomic survey sequence.  
 ACCESSION A0349458  
 VERSION A0349458.1 GI:4174354  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Bakayofci; Metzcoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homidae; Homo.  
 1 (bases 1 to 722)  
 Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and  
 Venter, J. C.  
 Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
 Map Building  
 Unpublished (1997)  
 Other GSSes: RPCI11-118J16.TV  
 Contact: Shaping Zhao, William Nierman, Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: hbeet@igf.org  
 Clones are derived from the human BAC library RPCI-11. For BAC  
 library availability, please contact Pieter de Jong  
 (pieterdejong.med.bu@ufla.edu). Clones may be purchased from  
 BACPAC Resources (<http://bacpac.med.bu@ufla.edu/ordering>) or from  
 Research Genetics ([info@resgen.com](http://info@resgen.com)). BAC end search page:  
[http://www.rti.org/rtdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.rti.org/rtdb/humgen/bac_end_search/bac_end_search.html)  
 Seq primer: SP6  
 Class: BAC ends.

FEATURES  
 source  
 Location/Qualifiers  
 1..722  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="GDB:7545159"  
 /db\_xref="taxon:9606"  
 /clone="RPCI-11-118J16"  
 /sex="Male"  
 /cell\_type="Lymphocytes"  
 /clone\_1fb="RPCI-11"  
 /note="Vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
 RPCI11 Human Male BAC Library"

ORIGIN  
 Query Match 1.4% Score 45; DB 9; Length 722;  
 Best Local Similarity 100.0%; Pred. No. 66-10;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy      3071 CAAGATTGGCACTGCAGCTCGGCGCAACAGCAAGACT 3115
      |||
      516 CAAGATTGGCACTGCAGCTCGGCGCAACAGCAAGACT 560

RESULT 96
AG009127
LOCUS      AG009127      726 bp      DNA      linear      GSS 16-FEB-2005
DEFINITION Homo sapiens genomic DNA, 21q region, clone: T172XN, genomic survey
sequence.
ACCESSION  AG009127 AG003036
VERSION     AG009127.1 GI:3289113
KEYWORDS   GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo.

REFERENCE
  1 Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
    Homo sapiens genomic DNA, chromosome 21q
    Published Only in Database (1998)
REFERENCE
  2 (bases 1 to 726)
  Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
    Direct Submision
    Submitted (06-JUN-1998) Masahira Hattori, RIKEN Genomic Sciences
    Center, RIKEN Yokohama Institute, Yokohama Research Promotion
    Division, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa,
    230-0045, Japan (E-mail:hattori@gsc.riken.jp, Tel:81-45-503-9111,
    Fax:81-45-503-9113)
    On Feb 5, 1999 this sequence version replaced gi:265425.
COMMENT     AG003036: Submitted (06-Dec-1997).
FEATURES
  source
    1..726
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="21"
    /map="21q"
    /clone="T172XN"

ORIGIN
Query Match      1.4%; Score 45; DB 10; Length 726;
Best Local Similarity 100.0%; Pred. No. 6e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2906 TGAAGCCAGAGATTGAGACCAAGCTGGCCAAATAGCGAAACC 2950
      |||
      179 TGAAGCCAGAGATTGAGACCAAGCTGGCCAAATAGCGAAACC 223

RESULT 97
AG009138
LOCUS      AG009138      731 bp      DNA      linear      GSS 16-FEB-2005
DEFINITION Homo sapiens genomic DNA, 21q region, clone: T172XN, genomic survey
sequence.
ACCESSION  AG009138 AG003047
VERSION     AG009138.1 GI:3289124
KEYWORDS   GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo.

REFERENCE
  1 Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
    Homo sapiens genomic DNA, chromosome 21q
    Published Only in Database (1998)
REFERENCE
  2 (bases 1 to 731)
  Hattori,M., Ishii,K., Toyoda,A., Shiba,T. and Sakaki,Y.
    Direct Submision
    Submitted (06-JUN-1998) Masahira Hattori, RIKEN Genomic Sciences
    Center, RIKEN Yokohama Institute, Yokohama Research Promotion

```

```

Division; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa,
230-0045, Japan (E-mail:hattori@gsc.riken.jp, Tel:81-45-503-9111,
Fax:81-45-503-9113)
On Feb 5, 1999 this sequence version replaced gi:265436.
COMMENT     AG003047: Submitted (06-Dec-1997).
FEATURES
  source
    1..731
    /organism="Homo sapiens"
    /mol_type="genomic DNA"
    /db_xref="taxon:9606"
    /chromosome="21"
    /map="21q"
    /clone="T172XN"

ORIGIN
Query Match      1.4%; Score 45; DB 10; Length 731;
Best Local Similarity 100.0%; Pred. No. 6e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2906 TGAAGCCAGAGATTGAGACCAAGCTGGCCAAATAGCGAAACC 2950
      |||
      381 TGAAGCCAGAGATTGAGACCAAGCTGGCCAAATAGCGAAACC 425

RESULT 99
BX457023
LOCUS      BX457023      754 bp      mRNA      linear      EST 06-MAY-2004

```

DEFINITION BX457023 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP005YD12  
5-PRIME, mRNA sequence.  
ACCESSION BX457023  
VERSION BX457023.2 GI:47069656  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 754)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On May 22, 2003 this sequence version replaced gi:31032820.  
Contact: Genoscope - Centre National de Sequencage  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
was not normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
This sequence belongs to sequence cluster 5067.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?c=CS0CAP005DB06Q1&c=5067.r.  
Location/Qualifiers  
1. 754  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0CAP005YD12"  
/tissue\_type="THYMUS"  
/clone\_lib="Homo sapiens THYMUS"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
Library was not normalized."

ORIGIN  
Query Match 1.4%; Score 45; DB 5; Length 754;  
Best Local Similarity 100.0%; Pred. No. 6e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGCCACTGCACCTCCAGCCTGGGCAACAGACAAGACTCTGTCTC 3122  
|||||  
622 GTGCCACTGCACCTCCAGCCTGGGCAACAGACAAGACTCTGTCTC 666

Db

RESULT 100  
LOCUS BX411269/3 768 bp mRNA linear EST 13-FEB-2001  
DEFINITION 602371333P1 NIH\_MGC\_93 Homo sapiens cDNA clone IMAGE:447632 5',  
mRNA sequence.  
ACCESSION BX411269  
VERSION BX411269.1 GI:12769865  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 768)  
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgapdb-r@mail.nih.gov  
Tissue Procurement: ATCC  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
http://image.llnl.gov  
Plate: LLAM10312 row: a column: 09  
High quality sequence atop: 700.  
Location/Qualifiers  
1. 768  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:447632"  
/tissue\_type="transitional cell papilloma, cell line"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_lib="NIH MGC 93"  
/note="Organ: bladder; Vector: pCMV-SPORT6; Site 1: NotI;  
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.  
Average insert size 1.7 kb. Library enriched for  
full-length clones and constructed by Life Technologies.  
Note: this is a NIH\_MGC library."

ORIGIN  
Query Match 1.4%; Score 45; DB 2; Length 768;  
Best Local Similarity 100.0%; Pred. No. 6e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCATCCTGAGCCGAGAGTTCGAGACAGCCTGCCACA 2939  
|||||  
Db 219 GGTGATCATCCTGAGCCGAGAGTTCGAGACAGCCTGCCACA 175

RESULT 101  
LOCUS BX411269/3 761 bp mRNA linear EST 03-MAY-2004  
DEFINITION BX411269 Homo sapiens FETAL BRAIN Homo sapiens cDNA clone  
CS0DP021YC03 3-PRIME, mRNA sequence.  
ACCESSION BX411269  
VERSION BX411269.2 GI:46955683  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
REFERENCE 1 (bases 1 to 761)  
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On May 15, 2003 this sequence version replaced gi:30767155.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
was not normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
This sequence belongs to sequence cluster 8085.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna?c=CS0BA10112B10\_CS01034\_2&c=8085.r  
Location/Qualifiers  
1. 761  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DP021YC03"  
/tissue\_type="FETAL BRAIN"  
/dev\_stage="fetal"  
/clone\_lib="Homo sapiens FETAL BRAIN"  
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA  
was primed with a NotI-oligo(dT) primer. Five prime end

## ORIGIN

enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

Query Match 1.4%; Score 45; DB 5; Length 781;  
Best Local Similarity 100.0%; Pred. No. 6e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2898 GGATCAGCTGAGGCGAGAGTTGAGACCGCTGCGCAACATAG 2342  
DB 723 GGATCAGCTGAGGCGAGAGTTGAGACCGCTGCGCAACATAG 679

## RESULT 102

CD242479 830 bp mRNA linear EST 22-MAY-2003  
LOCUS AGENCOURT 14120572 NIH MGC 179 Homo sapiens cDNA clone  
DEFINITION IMAGE:30384600 5', mRNA sequence.

CD242479  
VERSION CD242479.1 GI:31002943  
KEYWORDS EST.

ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

REFERENCE 1 (bases 1 to 830)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov  
Tissue Procurement: Dr. Michael Brownstein  
cDNA Library Preparation: Invitrogen Corp  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: NDAM450 row: 5 column: 01  
High quality sequence stop: 493.

FEATURES  
source 1. 830  
Location/Qualifiers

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:30384600"  
/tissue\_type="Pituitary"  
/lab\_host="DH10B-Ton A (T1 and T5 phage resistances)"  
/clone\_lib="NIH\_MGC\_179"  
/note="Organ: Brain; Vector: pCMV-SPORT6.1; Site 1: EcoRV (destroyed); Site 2: NotI; Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.1 kb. Library was constructed by (Invitrogen). Note: this is a NIH\_MGC library."

## ORIGIN

Query Match 1.4%; Score 45; DB 6; Length 830;  
Best Local Similarity 100.0%; Pred. No. 6e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2898 GTGATCAGCTGAGGCGAGAGTTGAGACCGCTGCGCAACAT 2940  
DB 436 GTGATCAGCTGAGGCGAGAGTTGAGACCGCTGCGCAACAT 540

RESULT 103  
LOCUS BU959380 856 bp mRNA linear EST 21-OCT-2002  
DEFINITION AGENCOURT 10622545 NIH MGC 127 Homo sapiens cDNA clone  
IMAGE:6737609 5', mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM

BU959380  
BU959380.1 GI:24188952  
EST.  
Homo sapiens (human)

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo  
1 (bases 1 to 856)  
NIH-MGC http://mgi.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov  
Tissue Procurement: NCI  
cDNA Library Preparation: Michael Brownstein Laboratory  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: LNCM3067 row: 0 column: 16  
High quality sequence stop: 431.

FEATURES  
source

1. 856  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6737609"  
/tissue\_type="mixed (pool of 40 RNAs)"  
/lab\_host="DH10B (T1-phage-resistant)"  
/clone\_lib="NIH\_MGC\_127"  
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccatggcc); Site 2: SfiI (ggccctggcc); Double-stranded cDNA was prepared from a pool of 40 cell line polyA+ RNAs (bladder - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows:  
5'-AAGCAGTGTATCAACGACGAGTGGCATTACGCGCGG-3' and  
5'-ATTCTAGAGCGGCGCGGACATG-dT(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the 1-2 kb size fraction (other fractions present in NIH\_MGC\_126 and NIH\_MGC\_128). Library created in the laboratory of T. Ueda, M.D., Ph.D. (NIH, NIH). Note: this is a NIH\_MGC library."

## ORIGIN

Query Match 1.4%; Score 45; DB 5; Length 856;  
Best Local Similarity 100.0%; Pred. No. 6e-10;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2888 TGAGCAGTGTGATCAGCTGAGGCGAGAGTTGAGACCGACTG 2932  
DB 203 TGAGCAGTGTGATCAGCTGAGGCGAGAGTTGAGACCGACTG 159

RESULT 104  
LOCUS B2771376 874 bp DNA linear GSS 13-MAR-2003  
DEFINITION m6876c11.g10 HPOSMID005 Homo sapiens genomic, genomic survey  
sequence.

ACCESSION B2771376  
VERSION B2771376  
KEYWORDS GSS.  
SOURCE B2771376.1 GI:28945045

ORGANISM Homo sapiens (human)  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.

```

REFERENCE      1 (bases 1 to 874)
AUTHORS        Cook,L., Delehaunt,K., Fewell,G., Fulton,L., Magrini,V.,
               Maris,E., Miner,T., Nash,W., Williams,D. and Wilson,R.K.
TITLE          Homo sapiens Fosmid End Reads
JOURNAL        Unpublished (2003)
COMMENT        Contact: Richard K. Wilson
               Genome Sequencing Center
               Washington University School of Medicine
               Email: submissions@wustl.edu
               Plate: mc876 row: c column: 11
               Class: fosmid ends
               High quality sequence start: 11
               High quality sequence stop: 518.
FEATURES       source
               1..874
               /organism="Homo sapiens"
               /mol_type="genomic DNA"
               /db_xref="taxon:9606"
               /clone_1ib="HPOSMD1005"
               /note="Vector: pcc01fos; Site 1: Eco721; Human whole
               genome fosmid library was prepared at Washington
               University Genome Sequencing Center. DNA was sheared for
               blunt-ended ligation into pcc01fos inducible vector. DNA
               was ordered from Coriell Cell Repository's DNA
               polymorphism discovery resource."
ORIGIN
Query Match      1.4%; Score 45; DB 9; Length 874;
Best Local Similarity 100.0%; Pred. No. 6e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTCACCTGCACCTGCAGCGGCGACAGACGACTCT 3117
DB      137 AGATTGTCACCTGCACCTGCAGCGGCGACAGACGACTCT 93

RESULT 105
CF596843/c      948 bp mRNA linear EST 26-SRP-2003
LOCUS           CF596843
DEFINITION      AGENCOURT.15668731 NICHD_Hs_Ov1 Homo sapiens cDNA clone
IMAGE:30705007 5', mRNA sequence.
ACCESSION      CF596843
VERSION        CF596843.1 GI:36353710
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 948)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Daniela S. Gerhard, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Gregory F. Erickson, Ph.D.
cDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: NDCM254 row: p column: 08
High quality sequence stop: 432.
FEATURES       source
               1..948
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:30705007"

```

```

/tissue_type="Ovary"
/lab_host="DH10B (T1 phage-resistant)"
/clone_1ib="NICHD_Hs_Ov1"
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);
Site 2: SfiI (ggccctctggcc); Library is oligo-dT primed
and directionally cloned. Granulosa lutein cells aspirated
from preovulatory follicles of normal cycling women
undergoing ovulation induction for infertility due to male
factor and normal donors. The cells were from follicles
stimulated with lupron, FSH and hCG. 5' and 3' adaptors
were used in cloning as follows: 5' adaptor sequence:
5'-CAAGCGCATTTATGACC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCCGCGCGCCGACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size 2.23
kb (range 1.0-4.5 kb). 14/15 colonies contained inserts by
PCR. This library was enriched for full-length clones and
was constructed by Clontech Laboratories (Palo Alto, CA)."
ORIGIN
Query Match      1.4%; Score 45; DB 6; Length 948;
Best Local Similarity 100.0%; Pred. No. 6e-10;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2895 GGTGATCACCCTGAGCCGAGGATTGAGACGAGCTGGCCACA 2939
DB      440 GGTGATCACCCTGAGCCGAGGATTGAGACGAGCTGGCCACA 396

RESULT 106
BQ896885/c      988 bp mRNA linear EST 16-AUG-2002
LOCUS           BQ896885
DEFINITION      AGENCOURT.8073985 NIH_MGC_102 Homo sapiens cDNA clone IMAGE:6087057
IMAGE:6087057 5', mRNA sequence.
ACCESSION      BQ896885
VERSION        BQ896885.1 GI:22288899
KEYWORDS       EST.
SOURCE         Homo sapiens (human)
ORGANISM       Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 988)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM2323 row: a column: 10
High quality sequence stop: 595.
FEATURES       source
               1..988
               /organism="Homo sapiens"
               /mol_type="mRNA"
               /db_xref="taxon:9606"
               /clone="IMAGE:6087057"
               /tissue_type="epidermoid carcinoma, cell line"
               /lab_host="DH10B (phage-resistant)"
               /clone_1ib="NIH_MGC_102"
               /note="Organ: salivary gland; Vector: pOTB7; Site 1: XhoI;
               Site 2: EcoRI; cDNA made by oligo-dT priming.
               Directionally cloned into EcoRI/XhoI sites using the
               following 5' adaptor: GGACGAG(G). Library constructed
               by Ling Hong in the laboratory of Gerald M. Rubin
               (University of California, Berkeley) using ZAP-cDNA
               synthesis kit (Stratagene) and Superscript II RT (Life
               Technologies). Note: this is a NIH_MGC Library."

```

ORIGIN

Query Match 1.4%; Score 45; DB 5; Length 988;  
 Best Local Similarity 100.0%; Pred. No. 6e-10;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACCCTGAGCCAGAGGTTGAGACCAAGCCTGGCCACACA 2939  
 53 GGTGATCACCCTGAGCCAGAGGTTGAGACCAAGCCTGGCCACACA 9

Db

RESULT 107  
 BX404721 1123 bp mRNA linear EST 01-MAY-2004  
 LOCUS BX404721 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP005YD12  
 DEFINITION 5-PRIME, mRNA sequence.  
 ACCESSION BX404721  
 VERSION BX404721.2 GI:46925367  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homidae; Homo.  
 1 (bases 1 to 1123)  
 L1.W.B., Gruber,C., Jessee,J. and Polayes,D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On May 13, 2003 this sequence version replaced gi:30639134.  
 COMMENT Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: sequef@genoscope.cns.fr; Web: www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
 was not normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen.  
 This sequence belongs to sequence cluster 5067.r  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna?c=CS0AMW132E04QPlc=5067.r.  
 FEATURES  
 source location/Qualifiers  
 1..1123  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0CAP005YD12"  
 /issue\_type="THYMUS"  
 /clone\_1ib="Homo sapiens THYMUS"  
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
 with a NotI-oligo(dT) primer. Five prime end enriched,  
 double-strand cDNA was digested with Not I and cloned into  
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
 Library was not normalized."

ORIGIN

Query Match 1.4%; Score 45; DB 5; Length 1123;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-10;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GGGCAGCTGACCTCCAGCTGGGCAACAGAGCAAGACTGTGCTC 3122  
 622 GGGCAGCTGACCTCCAGCTGGGCAACAGAGCAAGACTGTGCTC 666

Db

RESULT 108  
 BG959135 203 bp mRNA linear EST 12-JUN-2001  
 LOCUS BG959135  
 DEFINITION PM4-CT0806-180301-003-g01 CT0806 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BG959135  
 VERSION BG959135.1 GI:14377306  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homidae; Homo.  
 1 (bases 1 to 203)  
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,  
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,  
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and  
 Simpson,A.J.J.  
 Shotgun sequencing of the human transcriptome with ORF expressed  
 sequence tags  
 Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 10737800  
 CONTACT: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: asimpson@ludwig.org.br  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
 (http://www.ludwig.org.br/scripts/gethtml2.pl?l=PM4kt2-PM4-CT0806-  
 180301-003-g01kt3=2001-03-18kt4=1)  
 Seq primer: puc 18 forward  
 High quality sequence start: 17  
 High quality sequence stop: 127.  
 FEATURES  
 source location/Qualifiers  
 1..203  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_1ib="CT0806"  
 /note="Organ: colon; Vector: puc18; Site 1: SmaI; Site 2:  
 SmaI; A mini-library was made by cloning products derived  
 from ONESTEPS PCR (U.S. Letters Patent application No.  
 196,716 - Ludwig Institute for Cancer Research) profiles  
 into the puc 18 vector. Reverse transcription of tissue  
 mRNA and cDNA amplification were performed under low  
 stringency conditions."

ORIGIN

Query Match 1.4%; Score 44; DB 2; Length 203;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2897 TGGATCACCCTGAGCCAGAGGTTGAGACCAAGCCTGGCCACAT 2940  
 61 TGGATCACCCTGAGCCAGAGGTTGAGACCAAGCCTGGCCACAT 104

Db

RESULT 109  
 BG010132/c 219 bp mRNA linear EST 24-JUN-2001  
 LOCUS BG010132  
 DEFINITION MR3-GN0185-041200-015-f11 GN0185 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BG010132  
 VERSION BG010132.1 GI:12457017  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homidae; Homo.  
 1 (bases 1 to 219)  
 Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R.,  
 Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Costa,F.F.,  
 Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H.,  
 Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V.,  
 O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and



TITLE  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags

JOURNAL  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

COMMENT  
10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the PAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=MR3&c2=MR3-GN0185-  
041200-015-Flt&c3=2000-12-04&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 10  
High quality sequence stop: 219.  
Location/Qualifiers

FEATURES  
source  
1..219  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="GN0185"  
/note="Organ: placenta\_normal; Vector: puc18; Site: 1:  
SmaI; Site 2: SmaI; A mini-library was made by cloning  
products derived from ORESTES PCR (U.S. Letters Patent  
Application No. 196,716 - Ludwig Institute for Cancer  
Research) profiles into the pUC 18 vector. Reverse  
transcription of tissue mRNA and cDNA amplification were  
performed under low stringency conditions."

ORIGIN  
Query Match 1.4%; Score 44; DB 2; Length 219;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2897 TGGATCACCCTGAGCCAGGAGTTGAGACACAGCCCTGGCCACAT 2940  
218 TGGATCACCCTGAGCCAGGAGTTGAGACACAGCCCTGGCCACAT 175

Db

RESULT 110  
BF893386 281 bp mRNA linear EST 18-JAN-2001  
LOCUS QV3-MT0129-111100-427-h11 MT0129 Homo sapiens cDNA, mRNA sequence.  
DEFINITION BF893386  
ACCESSION BF893386  
VERSION BF893386.1 GI:12284845  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 281)  
Dias Neto, B., Garcia Correa, R., Verjovski-Almeida, S., Brijones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V.,  
O'Hare, M.U., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.G.  
Shotgun sequencing of the human transcriptome with ORF expressed  
sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
10737800  
Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil

Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the PAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=QV3&c2=QV3-MT0129-  
111100-427-h11&c3=2000-11-11&t4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 5  
High quality sequence stop: 281.  
Location/Qualifiers

FEATURES  
source  
1..281  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="MT0129"  
/note="Organ: marrow; Vector: puc18; Site: 1:  
SmaI; A mini-library was made by cloning products derived  
from ORESTES PCR (U.S. Letters Patent Application No.  
196,716 - Ludwig Institute for Cancer Research) profiles  
into the pUC 18 vector. Reverse transcription of tissue  
mRNA and cDNA amplification were performed under low  
stringency conditions."

ORIGIN  
Query Match 1.4%; Score 44; DB 2; Length 281;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 GGGCCATGCACTCCAGGCTGGGCAACAGAGCAAGACTCTGTCT 3121  
67 GTGCCATGCACTCCAGGCTGGGCAACAGAGCAAGACTCTGTCT 110

Db

RESULT 111  
BH609712 281 bp DNA linear GSS 18-DEC-2001  
LOCUS HTV18B03 Suprt HTV-1 in vivo integration lines Homo sapiens genomic  
DEFINITION HTV18B03, genomic survey sequence.  
ACCESSION BH609712  
VERSION BH609712.1 GI:17922321  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 281)  
Schroeder A.R.W., Shinn, P., Chen, H., Berry, C., Ecker, J.R. and  
Bushman, F.  
Favored Sites for HIV-1 Integration in the Human Genome  
Unpublished (2002)  
Contact: Frederic Bushman  
Salk Institute Infectious Disease Laboratory  
The Salk Institute for Biological Studies  
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
Tel: 858 453 4100 x1630  
Fax: 858 554 0341  
Email: bushman@salk.edu  
Class: PCR with specific primers.  
Location/Qualifiers

FEATURES  
source  
1..281  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="HTV18B03"  
/clone\_lib="Suprt1"  
/note="A human T-cell line (Suprt1) was infected with an  
HIV-based vector. DNA was isolated and cleaved with an  
restriction enzymes; linkers were ligated onto the cleaved  
DNA and DNAs were amplified using one primer that bound to

the linker DNA and one that bound to the HIV cDNA.  
Functions between integrated HIV proviruses and cellular  
DNA were cloned and sequenced."

## ORIGIN

Query Match 1.4%; Score 44; DB 9; Length 281;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3078 GTGCCACTGCACCTCCAGCTGGGCAACAGACAGACTCTGTCT 3121  
148 GTGCCACTGCACCTCCAGCTGGGCAACAGACAGACTCTGTCT 191

RESULT 112  
AA493894/C LOCUS 294 bp mRNA linear EST 20-AUG-1997  
DEFINITION nt07b12.s1 NCI CGAP Thy1 Homo sapiens cDNA clone IMAGE:943583  
similar to contains Alu repetitive element; contains element PTR5  
repetitive element ;, mRNA sequence.

ACCESSION AA493894 GI:2223735  
VERSION AA493894.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 294)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.  
Emmert-Buck, M.D., Ph.D.  
cDNA Library Preparation: David B. Krizman, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
[www.bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)  
Insert Length: 679 Std Error: 0.00  
Seq primer: -40m3 fwd. RT from Amersham  
High quality sequence stop: 256.  
Location/Qualifiers

FEATURES  
source  
1..294  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:943583"  
/tissue\_type="thyroid"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Thy1"  
/note="Vector: PAMF10; mRNA made from invasive thyroid  
tumor. cDNA made by oligo-dT priming. Non-directionally  
cloned. Size selected on agarose gel, average insert size  
600 bp. Reference: Krizman et al. (1996) Cancer Research  
56:5380-5383."

## ORIGIN

Query Match 1.4%; Score 44; DB 1; Length 294;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3079 TGCCACTGCACCTCCAGCTGGGCAACAGACAGACTCTGTCTC 3122  
74 TGCCACTGCACCTCCAGCTGGGCAACAGACAGACTCTGTCTC 31

RESULT 113  
A1382825/C

LOCUS A1382825 301 bp mRNA linear EST 18-MAR-1999  
DEFINITION ta72h10.x1 Soares total\_fetus Nb2HP8 9w Homo sapiens cDNA clone  
IMAGE:2049667 3' similar to contains Alu repetitive  
element; contains element MSRI repetitive element ;, mRNA sequence.

ACCESSION A1382825  
VERSION A1382825.1 GI:4195606  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 301)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgapbs-remail.nih.gov](mailto:cgapbs-remail.nih.gov)  
This clone is available royalty-free through LLNL; contact the  
IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
Insert Length: 923 Std Error: 0.00  
Seq primer: -40UP from Gldco  
High quality sequence stop: 287.  
Location/Qualifiers

FEATURES  
source  
1..301  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2049667"  
/dev\_stage="8-9 weeks"  
/lab\_host="DH10B"  
/clone\_lib="Soares total\_fetus Nb2HP8 9w"  
/note="Vector: pT73D-Pac (Pharmacia) with a modified  
polylinker; Site\_1: Not I; Site\_2: Eco RI; 1st strand cDNA  
was prepared from mRNA obtained from pooled 8-9 week  
(total) fetus material with a Not I - oligo(dT) primer [5'  
TGTTACCAATCTGAAGTGGGCGCGCTTAATTTTATTTT 3']  
Double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Not I and cloned into the Not I  
and Eco RI sites of the modified pT73 vector. Library  
went through one round of normalization, and was  
constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 1.4%; Score 44; DB 1; Length 301;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3073 AGATTGTCACCTGCACCTCCAGCTGGGCAACAGACAGACTC 3116  
Db 94 AGATTGTCACCTGCACCTCCAGCTGGGCAACAGACAGACTC 51

RESULT 114  
BF879334 LOCUS 301 bp mRNA linear EST 17-JAN-2001  
DEFINITION IIS-ET0111-231100-337-C04 ET0111 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF879334  
VERSION BF879334.1 GI:12269464  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 301)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,  
Nagai, M.A., da Silva, W. Jr., Zagó, M.A., Bordin, S., Costa, F.F.,  
Brunstein, G.H., Carvalho, A.F., Matukuma, A., Bata, G.S., Simpson, D.H.,  
Brunstein, A., deoliveira, P.S., Bucher, P., Jongeneel, C.V.,  
O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and  
Simpson, A.J.

**TITLE** Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

**PUBMED** 10737800

**COMMENT** Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=IL3&c2=IL3-UT0111-231100-337-C04&c3=2000-11-23&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 301.  
Location/Qualifiers  
1..301  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="ET0111"  
/note="Organ: Lung tumor; Vector: puc18; Site 1: Sma1; Site 2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

**ORIGIN**

Query Match 1.4%; Score 44; DB 2; Length 301;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 3073 AATATTGCGCAGCTCCAGCTGGGCAACAGAGACTGCTC 3116  
|||||  
183 AGATTGTGCGCAGCTCCAGCTGGGCAACAGAGACTGCTC 226

**Db** 183 AGATTGTGCGCAGCTCCAGCTGGGCAACAGAGACTGCTC 226

**RESULT 115**  
BF914587/c 325 bp mRNA linear EST 18-JAN-2001  
DEFINITION IL3-UT0114-011200-362-C02 UT0114 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BF914587  
VERSION BF914587.1 GI:12306045  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo.  
1 (bases 1 to 325)  
Dias Neto,B., Garcia Correa,R., Verjowski-Almeida,S., Britones,M.R., Nagai,M.A., da Silva,W. Jr., Zago,M.A., Bordin,S., Coeata,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brumstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

**JOURNAL** Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

**PUBMED** 10737800

**COMMENT** Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=IL3&c2=IL3-UT0114-011200-362-C02&c3=2000-12-01&t4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 325.  
Location/Qualifiers  
1..325  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="UT0114"  
/note="Organ: uterus tumor; Vector: puc18; Site 1: Sma1; Site 2: Sma1; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

**ORIGIN**

Query Match 1.4%; Score 44; DB 2; Length 325;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 3079 TCCCACTGCACTCCAGCTGGGCAACAGAGACTGCTGCTC 3122  
|||||  
224 TCCCACTGCACTCCAGCTGGGCAACAGAGACTGCTGCTC 181

**Db** 224 TCCCACTGCACTCCAGCTGGGCAACAGAGACTGCTGCTC 181

**RESULT 116**  
A0059204/c 337 bp DNA linear GSS 31-JUL-1998  
LOCUS A0059204.1  
DEFINITION CIT-HSP-2349B18.TR CIT-HSP Homo sapiens genomic clone 2349B18, genomic survey sequence.  
ACCESSION A0059204  
VERSION A0059204.1 GI:3361541  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homnidae; Homo  
1 (bases 1 to 337)  
Adams,M.D., Rounalely,S.D., Zhao,S., Field,C.E., Baes,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.  
Use of a random BAC End sequence Database for Sequence-Ready Map Building (1998)  
Unpublished (1998)  
Other GSSes: CIT-HSP-2349B18.TF  
Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mcdadam@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC clones are available from Research Genetics (info@resgen.com). BAC  
http://www.tigr.org/cdb/hungen/bac\_end\_search/bac\_end\_search.html.  
Seq primer: M13 Reverse  
Clas: BAC ends.  
Location/Qualifiers  
1..337  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="2349B18"  
/sex="Male"  
/cell\_type="Sperm"

**FEATURES**  
source

## ORIGIN

/clone.lib="CIT-HSP"  
/note="Vector: pBelOAC11; Site\_1: HindIII; Site\_2:  
HindIII"

Query Match 1.4%; Score 44; DB 9; Length 337;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGTGCGCACTGCATCCAGCTGGGCAACAGCAAGCACTC 3116  
Db 192 AGATTGTGCGCACTGCATCCAGCTGGGCAACAGCAAGCACTC 149

## RESULT 117

N63149/c 338 bp mRNA linear EST 01-MAR-1996  
LOCUS yz37e10.s1 Morton Fetal Cochlea Homo sapiens cDNA clone  
DEFINITION IMAGE:285258 3' similar to contains Alu repetitive element; mRNA  
sequence.

ACCESSION N63149  
VERSION N63149.1 GI:1210978  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 338)  
Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapell, B.,  
Chisoe, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,  
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N.,  
Martis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,  
Rohlfing, T., Schellberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,  
Trevasakis, B., Underwood, K., Wohldmann, P., Waterston, R., Wilson, R.  
and Marra, M.

TITLE Generation and analysis of 280,000 human expressed sequence tags  
JOURNAL Genome Res. 6 (9), 807-828 (1996)  
PUBMED 888549

## COMMENT

Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: ewatson@wustl.edu  
This clone is available royalty-free through LML; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: m13 -40 forward  
High quality sequence stop: 382.

## FEATURES

source 1.338  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:3892375"  
/db\_xref="taxon:9606"  
/clone="IMAGE:285258"  
/issue\_type="cochlea"  
/dev\_stage="16-22 week fetus"  
/lab\_host="SOLR cells (kanamycin resistant)"  
/note="Organ: ear; Vector: pBluescript SK-; Site 1: EcoRI;  
site 2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned  
unidirectionally. Primer: Oligo dT. Fetal cochlea, normal.  
37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb. Uni-ZAP  
XR Vector. Library constructed by N. Robertson, C. Morton.  
-5' adaptor sequence: 5' GAATTCCGACGAG 3' -3' adaptor  
sequence: 5' CTCGAGTTTCTTTTCTTTT 3' "

## ORIGIN

Query Match 1.4%; Score 44; DB 8; Length 338;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3079 TGCCACTGCACTCCAGCTGGGCAACAGCAAGCAAGCTGCTC 3122  
Db 44 TGCCACTGCACTCCAGCTGGGCAACAGCAAGCAAGCTGCTC 1

RESULT 118 352 bp DNA linear GSS 27-AUG-1998  
LOCUS A0092587/c HS\_3003\_B2\_A09\_MR CIT Approved Human Genomic Sperm Library D Homo  
DEFINITION sapiens genomic clone Plate=3003 Col=18 Row=B, genomic survey  
sequence.

ACCESSION A0092587  
VERSION A0092587  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 352)  
Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,  
Keller, A., Snaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and  
Hood, L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

CONTACT: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Sequence Tagged Connector  
Plate: 3003 row: B column: 18  
Class: BAC ends  
High quality sequence stop: 352.

## FEATURES

source 1.352  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="Plate=3003 Col=18 Row=B"  
/sex="male"  
/clone.lib="CIT Approved Human Genomic Sperm Library D"  
/note="Organ: sperm; Vector: pBelOAC11; BAC Clones in  
B-Col1 DH10B"

## ORIGIN

Query Match 1.4%; Score 44; DB 9; Length 352;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGTGCGCACTGCATCCAGCTGGGCAACAGCAAGCACTC 3116  
Db 165 AGATTGTGCGCACTGCATCCAGCTGGGCAACAGCAAGCACTC 122

## RESULT 119

A1367551/c 353 bp mRNA linear EST 15-FEB-1999  
LOCUS A1367551 GV9805.x1 NCI CGAP UC2 Homo sapiens cDNA clone IMAGE:1989728 3'  
DEFINITION similar to contains Alu repetitive element; contains element PTRS  
repetitive element; mRNA sequence.

ACCESSION A1367551  
VERSION A1367551.1 GI:4137296  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

REFERENCE 1 (bases 1 to 353)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 TITLE Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbbs-remail.nih.gov](mailto:cgapbbs-remail.nih.gov)  
 Tissue Procurement: Christopher Moskalkuk, M.D., Ph.D., Michael R.  
 Emmer-Buck, M.D., Ph.D.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/BLNL at:  
[www.bio.lnlnl.gov/bbrip/image/image.html](http://www.bio.lnlnl.gov/bbrip/image/image.html)  
 Insert length: 1947 Std Error: 0.00  
 Seq primer: -40UP from Glibco  
 High quality sequence stop: 352.  
 Location/Qualifiers  
 1..353  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1989728"  
 /tissue\_type="moderately-differentiated endometrial  
 adenocarcinoma, 3 pooled tumors"  
 /lab\_host="DH10B"  
 /clone\_11b="NCI-CGAP\_Ut2"  
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site: 1: SalI;  
 Site: 2: NotI; Cloned unidirectionally. Primer: Oligo dt.  
 Average insert size 1.85 kb. Life Technologies catalog #: 11539-012"

ORIGIN  
 Query Match 1.4%; Score 44; DB 1; Length 353;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TGCCACTGCATCTCCAGCTGGGCAACAGCAAGACTGTCTC 3122  
 |||||  
 DB 45 TGCCACTGCATCTCCAGCTGGGCAACAGCAAGACTGTCTC 2

RESULT 120  
 AA654781/C 365 bp mRNA linear EST 04-NOV-1997  
 LOCUS nt73g04.s1 NCI-CGAP\_P73 Homo sapiens CDNA clone IMAGE:1204182  
 DEFINITION similar to contains Alu repetitive element; contains element PTR5  
 MERRA repetitive element /, mRNA sequence.  
 ACCESSION AA654781  
 VERSION AA654781.1 GI:2550935  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 1 (bases 1 to 365)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgapbbs-remail.nih.gov](mailto:cgapbbs-remail.nih.gov)  
 Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai,  
 M.D., Michael Emmer-Buck, M.D., Ph.D.  
 CDNA Library Preparation: David B. Krizman, Ph.D.  
 CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/BLNL at:  
[www.bio.lnlnl.gov/bbrip/image/image.html](http://www.bio.lnlnl.gov/bbrip/image/image.html)  
 Seq primer: -40m13 fwd. BT from Amerham

FEATURES High quality sequence stop: 321.  
 source Location/Qualifiers  
 1..365  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1204182"  
 /sex="Male"  
 /dev stage="45 years old"  
 /lab\_host="DH10B"  
 /clone\_11b="NCI-CGAP\_P73"  
 /note="Vector: pAMP10; Site 1: NotI; Site 2: EcoRI; 1st  
 strand cDNA was primed with oligo(dT)17 on 50 ng of  
 DNase-treated, total cellular RNA obtained from  
 5,000-10,000 microdissected cells  
 histologically-determined to be fully malignant prostate  
 cancer cells. Double-stranded cDNA was ligated to EcoRI  
 adaptor, 5 cycles of PCR applied to the cDNA with an  
 adaptor-specific primer, and the resulting PCR product  
 subcloned into pAMP10 by the UDG-cloning method (Life  
 Technologies). Average insert size is 600 bp. NOTE: Not  
 directionally cloned. This library was constructed by  
 David Krizman."

ORIGIN  
 Query Match 1.4%; Score 44; DB 1; Length 365;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TGCCACTGCATCTCCAGCTGGGCAACAGCAAGACTGTCTC 3122  
 |||||  
 DB 74 TGCCACTGCATCTCCAGCTGGGCAACAGCAAGACTGTCTC 31

RESULT 121  
 A0061311 377 bp DNA linear GSS 31-JUL-1998  
 LOCUS CIT-HSP-2349M9.TP CIT-HSP Homo sapiens genomic clone 2349M9,  
 DEFINITION genomic survey sequence.  
 ACCESSION A0061311  
 VERSION A0061311.1 GI:3363223  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homnidae; Homo.  
 1 (bases 1 to 377)  
 Adams M.D., Rounale, S.D., Zhao S., Field, C.E., Bass S., Linher, K.,  
 Simon, M., and Venter, J.C.  
 Use of a random BAC End Sequence Database for Sequence-Ready Map  
 Building (1998)  
 Unpublished (1998)  
 Other GSS: CIT-HSP-2349M9.TR  
 CONTACT: Mark Adams  
 Department of Eukaryotic Genomics  
 The Institute for Genomic Research  
 9712 Medical Center Dr., Rockville, MD 20850, USA  
 Tel: 301 838 0200  
 Fax: 301 838 0208  
 Email: [madams@tigr.org](mailto:madams@tigr.org)  
 Clones are available from Research Genetics (info@resgen.com). BAC  
 end search page:  
[http://www.tigr.org/tcb/hungen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/tcb/hungen/bac_end_search/bac_end_search.html).  
 Seq primer: M13-21  
 Class: BAC ends.

FEATURES Location/Qualifiers  
 source 1..377  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="2349M9"

ORIGIN

/sex="Male"  
/cell\_type="Sperm"  
/clone\_lib="CIT-HSP"  
/note="Vector: pBelobAC11; Site\_1: HindIII; Site\_2: HindIII"

Query Match 1.4%; Score 44; DB 9; Length 377;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGACGAGTGATCAGCTGAGCCAGAGATTGAGACGAGCCTG 2932  
|||||  
209 GAGGACGAGTGATCAGCTGAGCCAGAGATTGAGACGAGCCTG 252

RESULT 122 378 bp mRNA linear EST 04-NOV-1997  
AA653916/c nt80a05.s1 NCI CGAP Pr3 Homo sapiens CDNA clone IMAGE:1204784  
LOCUS  
DEFINITION similar to contigins Alu repetitive element; contains element MSRI  
repetitive element; mRNA sequence.

ACCESSION AA653916 GI:2590070  
VERSION AA653916  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

REFERENCE 1 (bases 1 to 378)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)

JOURNAL  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-rcmail.nih.gov  
Tissue Procurement: W. Marston Linehan, M.D., Rodrigo Chuquai,  
M.D., Michael Emmert-Buck, M.D., Ph.D.

CDNA Library Preparation: David B. Krizman, Ph.D.  
CDNA Library Arrayed by: Genome Systems Inc., Greg Lennon, Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LIN ac:  
[www.bio.lnlnl.gov/bbrp/image/image.html](http://www.bio.lnlnl.gov/bbrp/image/image.html)  
Seq primer: 40m3 fwd. ET from Amerham.

FEATURES  
source Location/Qualifiers  
1..378

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1204784"  
/sex="Male"  
/dev\_stage="45 years old"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP Pr3"  
/note="Vector: PAMPI0; Site\_1: NotI; Site\_2: EcoRI; 1st  
strand CDNA was primed with oligo(dT)17 on 50 ng of  
DNase-treated, total cellular RNA obtained from  
5,000-10,000 microdissected cells  
histologically-determined to be fully malignant prostate  
cancer cells. Double-stranded CDNA was ligated to EcoRI  
adaptors, 5 cycles of PCR applied to the CDNA with an  
adaptor-specific primer, and the resulting PCR product  
subcloned into PAMPI0 by the UDC-cloning method (Life  
Technologies). Average insert size is 600 bp. NOTE: Not  
directionally cloned. This library was constructed by  
David Krizman."

## ORIGIN

Query Match 1.4%; Score 44; DB 1; Length 378;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
|||||  
Db 50 TGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 7

RESULT 123 381 bp DNA linear GSS 30-SEP-1998  
AQ239365/c CIT-HSP-2383D24.TR.1 CIT-HSP Homo sapiens genomic clone 2383D24,  
genomic survey sequence.

ACCESSION AQ239365  
VERSION AQ239365.1 GI:3671656  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

REFERENCE 1 (bases 1 to 381)  
Adams,M.D., Rounsley,S.D., Zhao,S., Base,S., Linher,K., Golden,K.,  
Berry,K., Granger,D., Sun,E., Wible,C., Shizuya,H., Simon,M. and  
Venter,J.C.  
Use of a random human BAC End Sequence Database for Sequence-Ready  
Map Building  
Unpublished (1998)

JOURNAL  
COMMENT Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208

Email: mda@msc.tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC  
end search page:  
[http://www.tigr.org/cdb/humgen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/cdb/humgen/bac_end_search/bac_end_search.html).  
Seq primer: M13 Reverse  
Class: BAC ends.

FEATURES  
source Location/Qualifiers  
1..381

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="2383D24"  
/sex="Male"  
/cell\_type="Sperm"  
/clone\_lib="CIT-HSP"  
/note="Vector: pBelobAC11; Site\_1: HindIII; Site\_2:  
HindIII"

## ORIGIN

Query Match 1.4%; Score 44; DB 9; Length 381;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCT 3121  
|||||  
Db 211 GTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCT 168

RESULT 124 382 bp mRNA linear EST 16-SEP-2002  
BUS61388 AGENCORRT\_10278700 NIH\_MGC\_82 Homo sapiens CDNA clone IMAGE:6592405  
DEFINITION 5', mRNA sequence.

ACCESSION BUS61388  
VERSION BUS61388.1 GI:22911684  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

REFERENCE 1 (bases 1 to 382)  
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
 Tissue Procurement: CLONTECH  
 cDNA Library Preparation: CLONTECH Laboratories, Inc.  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation (LNL)  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>  
 Plate: LNCM2815 row: 1 column: 13  
 High quality sequence stop: 381.  
 Location/Qualifiers

## FEATURES

1..382  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:6592405"  
 /lab\_host="DH10B (T1 phage-resistant)"  
 /clone\_1lb="NIH\_MGC\_82"  
 /note="Organ: testis; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggccgctcgcc); Site 2: SfiI (ggccatcgcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTAGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGAGCGGCCGACATG-dt(30)BR-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.35 kb (range 0.9-4.0 kb). 14/15 colonies inserted inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

## ORIGIN

Query Match 1.4%; Score 44; DB 5; Length 382;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 3122  
 |||||||  
 Db 233 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 276

RESULT 125  
 AA728990 389 bp mRNA linear EST 06-JAN-1998  
 LOCUS nw22906.s1 NCI\_CGAP GCBO Homo sapiens cDNA clone IMAGE:1241242 3'  
 DEFINITION similar to contains Alu repetitive element; contains 11.11 L1  
 repetitive element; mRNA sequence.

ACCESSION AA728990.1 GI:2750349  
 VERSION AA728990  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homidae; Homo.  
 1 (bases 1 to 389)  
 NCI\_CGAP <http://www.nci.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
 Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,  
 Ph.D., Gerald Marti, M.D.  
 cDNA Library Preparation: Louis M. Staudt, M.D., Ph.D.  
 cDNA Library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: [www.bio.llnl.gov/bbrp/image/image.html](http://www.bio.llnl.gov/bbrp/image/image.html)

REFERENCE  
 AUTHORS  
 TITLE  
 JOURNAL  
 COMMENT

Seq primer: -40m13 fwd. BT from Amerham  
 High quality sequence stop: 367.  
 Location/Qualifiers

## FEATURES

1..389  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1241242"  
 /tissue\_type="germinal center B-cells"  
 /lab\_host="SOLR (kanamycin resistant)"  
 /clone\_1lb="NCI\_CGAP\_GCBO"  
 /note="Organ: tonsil; Vector: Bluescript SK-; Site 1: EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer: Oligo dt. Germinal center B-cells library constructed by Dr. L. Staudt (NCI). 5' adaptor sequence: 5' GATTTCGACGAG 3' 3' adaptor sequence: 5' CTCAGTTTTTTTTTTTTTTT 3' Average insert size: 1.1 kb."

## ORIGIN

Query Match 1.4%; Score 44; DB 1; Length 389;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 3122  
 |||||||  
 Db 58 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 15

RESULT 126  
 H44630 409 bp mRNA linear EST 31-JUL-1995  
 LOCUS YP19G10.s1 Soares breast 3NBHst Homo sapiens cDNA clone  
 DEFINITION IMAGE:187938 3' similar to contains Alu repetitive element; mRNA  
 sequence.

ACCESSION H44630  
 VERSION H44630.1 GI:920682  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homidae; Homo.  
 1 (bases 1 to 409)  
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
 Parsons, J., Rifkin, B., Rohlfing, T., Soares, M., Tan, F.,  
 Trevaaskis, E., Waterston, R., Williamson, A., Wohldmann, P. and  
 Wilson, R.  
 The WashU-Merck EST Project  
 Unpublished (1995)  
 Contact: Wilson R.  
 Washington University School of Medicine  
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: [est@watson.wustl.edu](mailto:est@watson.wustl.edu)  
 Insert size: 1706  
 High quality sequence stops: 345  
 Source: IMAGE Consortium, LNL  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium ([info@image.llnl.gov](mailto:info@image.llnl.gov)) for further information.  
 Insert length: 1706 Std Error: 0.00  
 Seq primer: Promega -21m13  
 High quality sequence stop: 345.  
 Location/Qualifiers

## FEATURES

1..409  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:3818835"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:187938"  
 /sex="Female"  
 /dev\_stage="adult"



/lab host="DH10B (ampicillin resistant)"  
 /clone\_lib="Soares breast 3bHbstr"  
 /note="Organ: breast; Vector: pT73D (Pharmacia) with a  
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer [5'  
 TGTTCACATCTGAGTGGAGCGCGCCCTTTTCTTTTCTTTT 3'],  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of a modified pT73 vector (Pharmacia).  
 Library went through one round of normalization to a Cot =  
 20. Library constructed by Bento Soares and M.Fatima  
 Bonaldo."

## ORIGIN

Query Match 1.4%; Score 44; DB 8; Length 409;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACTGCACTCCAGCTCGGCAACAGAGCAAGACTC 3116  
 Db 314 AGATTGTGCACTGCACTCCAGCTCGGCAACAGAGCAAGACTC 357

RESULT 127  
 A0605062 412 bp DNA linear GSS 10-JUN-1999  
 LOCUS HS\_2119\_B1\_G05\_T7C CIT Approved Human Genomic Sperm Library D Homo  
 DEFINITION sapiens genomic clone Plate=2119 Col=9 Row=N, genomic survey  
 sequence.

ACCESSION A0605062  
 VERSION A0605062.1 GI:5065056  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 412)  
 Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holman,T.,  
 Kellier,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,W.D. and  
 Hood,L.  
 Sequence-tagged connectors: A sequence approach to mapping and  
 scanning the human genome  
 Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

TITLE  
 JOURNAL  
 PUBMED  
 COMMENT  
 Contact: Mahairas GG, Wallace JC, Hood L  
 High Throughput Sequencing Center  
 University of Washington  
 401 Queen Anne Avenue North, Seattle, WA 98109, USA  
 Tel: (206) 616-3618  
 Fax: (206) 616-3887  
 Email: jwallace@u.washington.edu  
 Clones may be purchased from Research Genetics (info@resgen.com).  
 BAC end Web Server: http://www.htsc.washington.edu  
 Plate: 2119 row: N column: 9  
 Seq primer: T7  
 Class: BAC ends  
 High quality sequence stop: 412.

## FEATURES

source  
 1. .412  
 /organism="Homo sapiens"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:9606"  
 /clone="Plate=2119 Col=9 Row=N"  
 /sex="male"  
 /clone\_1b="CIT Approved Human Genomic Sperm Library D"  
 /note="Organ: sperm; Vector: pBel0BAC11; BAC Clones in  
 E-Coli DH10B"

## ORIGIN

Query Match 1.4%; Score 44; DB 9; Length 412;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACTGCACTCCAGCTCGGCAACAGAGCAAGACTC 3116  
 Db 88 AGATTGTGCACTGCACTCCAGCTCGGCAACAGAGCAAGACTC 45

RESULT 128  
 CR769029/C 421 bp mRNA linear EST 23-SEP-2004  
 LOCUS DKFZp46811312\_r1 468 (synonym: phr1) Pongo pygmaeus cDNA clone  
 DEFINITION DKFZp46811312 5', mRNA sequence.  
 CR769029  
 ACCESSION CR769029  
 VERSION CR769029  
 KEYWORDS EST.  
 SOURCE Pongo pygmaeus (orangutan)  
 ORGANISM Pongo pygmaeus

REFERENCE 1 (bases 1 to 421)  
 Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,  
 Mewes,H.W., Weil,B., Amlid,C., Osanger,A., Fobo,G., Han,M. and  
 Wiemann,S.  
 Pongo pygmaeus mRNA (Poustka,A., Albert,R., Moosmayer,P., et al.)  
 Unpublished (2004)  
 CONTACT: MIPS

TITLE  
 JOURNAL  
 COMMENT  
 MIPS  
 Ingolstaedter Landstr.1, D-85764 Neuberg, Germany  
 This is the 5' sequence of the clone insert. Clone from S. Wiemann,  
 Molecular Genome Analysis, German Cancer Research Center (DKFZ);  
 Email s.wiemann@dkfz-heidelberg.de; any. please contact RZPD for  
 ordering:  
 http://www.rzpd.de/cgi-bin/products/cl.cgi?cloneid=DKFZp46811312  
 Further information about the clone and the sequencing project is  
 available at http://mips.gsf.de/projects/cdna/.

FEATURES  
 source  
 1. .421  
 /organism="Pongo pygmaeus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9600"  
 /clone="DKFZp46811312"  
 /issue\_type="heart"  
 /dev stage="adult"  
 /lab\_host="DH10B"  
 /clone\_1b="468 (synonym: phr1)"  
 /note="Vector: pSport1\_Sfi; Site\_1: SfiI; Site\_2: SfiIb"

## ORIGIN

Query Match 1.4%; Score 44; DB 7; Length 421;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACTGCACTCCAGCTCGGCAACAGAGCAAGACTC 3116  
 Db 145 AGATTGTGCACTGCACTCCAGCTCGGCAACAGAGCAAGACTC 102

## RESULT 129

N20066 426 bp mRNA linear EST 15-DEC-1995  
 LOCUS YX28904.s1 Soares melanocyte 2Nbh Homo sapiens cDNA clone  
 DEFINITION IMAGE:263094 3' similar to contains Alu repetitive element, mRNA  
 sequence.

ACCESSION N20066  
 VERSION N20066.1 GI:1124733  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 426)  
 Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,  
 Mammalia; Eutheria; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homiidae; Homo.  
 1 (bases 1 to 426)  
 Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M.,

**TITLE**  
Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M.,  
Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, F.,  
Trevaskis, E., Waterston, R., Williamson, A., Wohlmann, P. and  
Wilson, R.  
The Marsh-Werck EST Project  
Unpublished (1995)  
**JOURNAL**  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu

**FEATURES**  
source  
High quality sequence stops: 404  
Source: IMAGS Consortium, LNL.  
This clone is available royalty-free through LNL; contact the  
IMAGS Consortium (info@image.lnl.gov) for further information.  
Insert Length: 2002 Std Error: 0.00  
Seq primer: m13 -40 forward  
High quality sequence stop: 404.  
Location/Qualifiers  
1..426  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:3872736"  
/db\_xref="taxon:9606"  
/clone="IMAGS:263094"  
/sex="Male"  
/tissue\_type="melanocyte"  
/lab\_host="PH10B (ampicillin resistant)"  
/clone\_lib="Soares melanocyte 2NBMH"  
/note="Vector: pT73D (Pharmacia) with a modified  
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
was primed with a Not I - oligo(dT) primer [5'  
TGTTCACATCTGAGTGGAGGCGGCGAGTGTCTTTTCTTTT 3']  
double-stranded cDNA was size selected, ligated to Eco RI  
adapters (Pharmacia), digested with Not I and cloned into  
the Not I and Eco RI sites of a modified pT73 vector  
(Pharmacia). Library constructed by Bentho Soares and  
M. Fatima Bonaldi. RNA from normal foreskin melanocytes  
(FS374) was kindly provided by Dr. Anthony P. Albino."

## ORIGIN

**Query Match**  
Best Local Similarity 100.0%; Score 44; DB 8; Length 426;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 3073 AGATTGTGCGCATCTCCAGCTCGGCAACAGCAAGACTC 3116  
|||||  
Db 380 AGATTGTGCGCATCTCCAGCTCGGCAACAGCAAGACTC 423  
|||||

**RESULT 130**  
CR546695/c 435 bp mRNA linear EST 07-JUL-2004  
**LOCUS** DKFZP470L1324.1 470 (synonym: p11v1) Pongo pygmaeus cDNA clone  
**DEFINITION** DKFZP470L1324.5, mRNA sequence.  
**ACCESSION** CR546695  
**VERSION** CR546695.1 GI:49899216  
**KEYWORDS** EST.  
**SOURCE** Pongo pygmaeus (orangutan)  
**ORGANISM** Pongo pygmaeus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Pongo.  
1 (bases 1 to 435)  
Wamburt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., Osaenger, A.,  
Fodor, G., Han, M., and Wiemann, S.  
Pongo pygmaeus mRNA (Wamburt, R., Heubner, D., Mewes, H.W., et al.)  
Unpublished (2004)  
**TITLE**  
Contact: MIPS  
MIPS  
Ingolstaedter Landstr.1, D-85764 Neuberg, Germany  
This is the 5' sequence of the clone insert from S. Wiemann,

**FEATURES**  
source  
Molecular Genome Analysis, German Cancer Research Center (DKFZ);  
Email: s.wiemann@dkfz-heidelberg.de; sequenced by Agowa  
(Berlin/Germany) within the cDNA sequencing consortium of the  
German Genome Project. This clone (DKFZP470L1324) is available at  
the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum,  
Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email:  
clone@rzpd.de Further information about the clone and the  
sequencing project is available at  
http://mips.gsf.de/projects/cdna/  
location/Qualifiers  
1..435  
/organism="Pongo pygmaeus"  
/mol\_type="mRNA"  
/db\_xref="taxon:9600"  
/clone="DKFZP470L1324"  
/tissue\_type="liver"  
/dev\_stage="adult"  
/lab\_host="PH10B"  
/clone\_lib="470 (synonym: p11v1)"  
/note="Vector: pSport1\_Sfi, Site\_1: Sfi1A, Site\_2: Sfi1B"

## ORIGIN

**Query Match**  
Best Local Similarity 100.0%; Score 44; DB 7; Length 435;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY** 3079 TGGCAGTGCATCTCCAGCTCGGCAACAGCAAGACTCTGTC 3122  
|||||  
Db 178 TGGCAGTGCATCTCCAGCTCGGCAACAGCAAGACTCTGTC 135  
|||||

**RESULT 131**  
B60059/c 438 bp DNA linear GSS 21-JUN-1998  
**LOCUS** CIT-HSP-385015.TRB CIT-HSP Homo sapiens genomic clone 385015,  
**DEFINITION** genomic survey sequence.  
**ACCESSION** B60059  
**VERSION** B60059.1 GI:2614777  
**KEYWORDS** GSS.  
**SOURCE** Homo sapiens (human)  
**ORGANISM** Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 438)  
Adams, M.D., Rounsailey, S.D., Field, C.E., Baas, S., Linher, K.,  
Golden, K., Berry, K., Granger, D., Suh, E., Wible, C., Shizuya, H.,  
Simon, M., and Venter, J.C.  
Use of a random BAC End Sequence Database for Sequence-Ready Map  
Building  
Unpublished (1997)  
Other\_GSSs: CIT-HSP-385015.TRB  
**CONTACT** Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdadams@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC  
end search page:  
http://www.tigr.org/tcdb/hungen/bac\_end\_search/bac\_end\_search.html  
Seq primer: M13-21  
Classes: BAC ends.  
location/Qualifiers  
1..438  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:537896"  
/db\_xref="taxon:9606"  
/clone="385015"  
/sex="Male"  
/cell\_type="Sperm"

## FEATURES

source  
1..438  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:537896"  
/db\_xref="taxon:9606"  
/clone="385015"  
/sex="Male"  
/cell\_type="Sperm"

## ORIGIN

/clone.lib="CIT-HSP"  
/note="Vector: pBelOAC11; Site\_1: HindIII; Site\_2:  
HindIII"

Query Match 1.4%; Score 44; DB 9; Length 438;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGGCACTGCACTCCAGCTGGGCAAGAGCAAGACTC 3116  
Db 436 AGATTGGCACTGCACTCCAGCTGGGCAAGAGCAAGACTC 393

## RESULT 132

AL713064 447 bp mRNA linear EST 04-SEP-2003  
LOCUS DKFZp686C1395.F1.686 (synonym: hlcc3) Homo sapiens cDNA clone  
DEFINITION DKFZp686C1395.5, mRNA sequence.

AL713064  
VERSION AL713064.1 GI:19696420  
KEYWORDS EST.

SOURCE Homo sapiens (human)  
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homidae; Homo.  
1 (bases 1 to 447)

REFERENCE Bloecher, H., Boecher, M., Brandt, P., Mewes, W., Well, B. and  
Wiemann, S.  
EST (Bloecher, H., Boecher, M., Brandt, P., Mewes, H.W., Well, B. and  
Wiemann, S.)  
Unpublished (1999)

JOURNAL  
COMMENT  
Contact: MIPS

MIPS

Ingolstedter Landstr. 1, D-85764 Neuberg, Germany  
This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
sequenced by GPF (National Research Centre for Biotechnology Ltd.,  
Brunschweig/Germany) within the cDNA sequencing consortium of the  
German Genome Project.

No sequence available.

This clone (DKFZp686C1395) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES  
source  
Location/Qualifiers

1..447  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp686C1395"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone.lib="686 (synonym: hlcc3)"  
/note="Vector: pT1p1Ex2; Site\_1: SfiI; Site\_2: SfiIb;  
cDNA-collection"

## ORIGIN

Query Match 1.4%; Score 44; DB 1; Length 447;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3079 TGGCACTGCACTCCAGCTGGGCAAGAGCAAGACTCTGTC 3122  
Db 264 TGGCACTGCACTCCAGCTGGGCAAGAGCAAGACTCTGTC 307

## RESULT 133

AL469968 448 bp mRNA linear EST 14-APR-1999  
LOCUS U189C03.x1 Soares NSF P8.9W OT PA.P.S1 Homo sapiens cDNA clone  
DEFINITION IMAGE:2148676.3, similar to contains Alu repetitive element;; mRNA

sequence.

AL469968  
VERSION AL469968.1 GI:4332058

KEYWORDS

SOURCE Homo sapiens (human)  
ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
1 (bases 1 to 448)

REFERENCE NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)

CONTACT: Robert Strausberg, Ph.D.  
Email: cgaap@nci.nih.gov  
This clone is available royalty-free through LNL; contact the  
IMAGE Consortium (info@image.llnl.gov) for further information.

Insert length: 710 Std Error: 0.00  
Seq primer: -40UP from Gibco  
High quality sequence stop: 447.  
Location/Qualifiers

FEATURES  
source

1..448  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:2148676"  
/lab\_host="DH10B"  
/clone.lib="Soares NSF P8.9W OT PA.P.S1"  
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with  
a modified polylinker; Site\_1: Not I; Site\_2: Eco RI;  
Equal amounts of plasmid DNA from five normalized  
libraries were mixed, and six circles were made in vitro.  
Following HAP purification, this DNA was used as tracer in  
a subtractive hybridization reaction. The diver was  
PCR-amplified cDNAs from pools of 5,000 clones made from  
the same 5 libraries. The pools consisted of the following  
libraries and clones: Soares NBSP pool 1:  
309384-310919, 323208-325895 Soares NBSP pool 1:  
145032-147335, 147720-148103, 148872-149255, 15002 -  
150407, 151176-152327 Soares NBSP pool 1:  
758280-760583, 772104-774407 Soares NBSP pool 1:  
304776-306311, 320136-322823, 326280-326663 Soares NBSP  
pool 1: 723720-726407, 739080-740999 Subtraction by Bento  
Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 1.4%; Score 44; DB 1; Length 448;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGGCACTGCACTCCAGCTGGGCAAGAGCAAGACTC 3116  
Db 381 AGATTGGCACTGCACTCCAGCTGGGCAAGAGCAAGACTC 424

## RESULT 134

AQ705037 448 bp DNA linear GSS 07-JUL-1999  
LOCUS HS.5521.B2.H07.T7A.RBC1-11 Human Male BAC library Homo sapiens  
DEFINITION genomic clone Plate=1097 Col=14 Row=P, genomic survey sequence.

AQ705037  
VERSION AQ705037.1 GI:5414463

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homidae; Homo.  
1 (bases 1 to 448)

REFERENCE Mahares, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,  
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and  
Hood, L.

**TITLE**  
Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

**JOURNAL**  
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

**PMID**  
10449764

**COMMENT**  
Contact: Mahatras GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@dejong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (http://bacpac.med.buffalo.edu/ordering/bac.htm) or from Research Genetics (info@resgen.com). BAC end Web Server: http://www.htsc.washington.edu  
Plate: 1097 row: P column: 14  
Seq primer: T7  
Class: BAC ends  
High quality sequence stop: 448.

**FEATURES**  
source  
1..448  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="plate:1097 Col=14 Row=P"  
/sex="male"  
/clone\_lib="RPCI-11 Human Male BAC Library"  
/note="vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"

**ORIGIN**  
Query Match 1.4%; Score 44; DB 9; Length 448;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY**  
3079 TGGCACTGCATCTCCAGCTGGGCAAGAGCAAGACTCTGTCTC 3122  
|||||  
391 TGGCACTGCATCTCCAGCTGGGCAAGAGCAAGACTCTGTCTC 348

**Db**  
391 TGGCACTGCATCTCCAGCTGGGCAAGAGCAAGACTCTGTCTC 348

**RESULT 135**  
AUI57238 466 bp mRNA linear EST 05-AUG-2002  
AUI57238 PLACB1 Homo sapiens cdna clone PLACB1006883 3', mRNA  
sequence.

**ACCESSION**  
AUI57238  
VERSION  
AUI57238.1 GI:11018759  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens

**REFERENCE**  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 466)  
Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Ozawa, M., Nakamura, Y., Nagai, T., Sugano, S. and Isogai, T.)  
HRI human CDNA project (Ota, T., Nishikawa, T., Suzuki, Y., Ishii, S., Saito, K., Kawai, Y., Yamamoto, J., Wakamatsu, A., Ozawa, M., Nakamura, Y., Nagai, T., Sugano, S., Isogai, T.)  
Unpublished (2000)

**JOURNAL**  
Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 222-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp

**TITLE**  
HRI human CDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; CDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

**COMMENT**  
Contact: Helix Research Institute.

**FEATURES**  
source  
1..466  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1340508"  
/issue\_type="Placenta"  
/clone\_lib="PLACB1"  
/note="vector: PMB18SFL3"

**ORIGIN**  
Query Match 1.4%; Score 44; DB 1; Length 466;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

**QY**  
3079 TGGCACTGCATCTCCAGCTGGGCAAGAGCAAGACTCTGTCTC 3122  
|||||  
44 TGGCACTGCATCTCCAGCTGGGCAAGAGCAAGACTCTGTCTC 1

**Db**  
44 TGGCACTGCATCTCCAGCTGGGCAAGAGCAAGACTCTGTCTC 1

**RESULT 136**  
BX111733 466 bp mRNA linear EST 07-FEB-2003  
BX111733 NCI CGAP GCBI Homo sapiens cdna clone IMAGE1340508; mRNA sequence.  
BX111733  
IMAGE:1340508, mRNA sequence.

**ACCESSION**  
BX111733  
VERSION  
BX111733.1 GI:27837211  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens

**REFERENCE**  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 466)  
Edert, L., Hell, O., Hennig, S., Neubert, P., Partsch, E., Peters, M., Radcliff, U., Schneider, D. and Korn, B.  
Human Unigeneset - RZPD3  
Unpublished (2003)  
Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
RZPD; IMAGE998D13381.  
RZPD; I.M.A.G.E. CDNA Clone Collection;  
Human Unigeneset - RZPD3 (RZPD LIB No. 972)  
http://www.rzpd.de/CloneCards/cgi-bin/showLib.pl.cgi?response=11bNo=972 Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Heidenberg 6, D-14059 Berlin, Germany  
Tel: +49 30 32639 101  
Fax: +49 30 32639 111  
www.rzpd.de

**JOURNAL**  
This clone is available royalty-free from RZPD; contact RZPD (clone@rzpd.de) for further information. Seq primer: M13r, Primer sequence: TTTCACACAGAAACAGCTATGAC.  
Location/Qualifiers  
1..466  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE998D13381; IMAGE:1340508"  
/issue\_type="germlinal center B cell"  
/lab\_host="DH10B"  
/clone\_lib="NCI CGAP GCBI"  
/note="vector: pTZ19-pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand CDNA was prepared from human tonsillar cells enriched for germinal center B cells by flow sorting (CD20+, IgD-), provided by Dr. Louis M. Staudt (NCI), Dr. David Altman (NCI) and Dr. Gerald Marti (CBER). CDNA synthesis was

primed with a Not I - oligo(dT) primer  
 [5'-TGTTCACATCTGAGAGTGGAGCGGCCCTTCATTTTCTTTT-3'  
 ]. Double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization, and was  
 constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 1.4%; Score 44; DB 5; Length 466;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCGACCTCCAGCTGGGCAAGAGCAAGACTC 3116  
 DB 83 AGATTGTGCGACCTCCAGCTGGGCAAGAGCAAGACTC 126

RESULT 137  
 BE674703/c 466 bp mRNA linear EST 08-SEP-2000  
 LOCUS 7e94e04.x1 NCI\_CGAP\_CLL1 Homo sapiens cDNA clone IMAGE:3292830 3'  
 DEFINITION similar to contains Alu repetitive element; contains element MIR  
 repetitive element ; mRNA sequence.

ACCESSION BE674703 GI:10035325  
 VERSION BE674703  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homidae; Homo.

REFERENCE 1 (bases 1 to 466)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)

## JOURNAL

## COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: cgapbs-remail.nih.gov  
 Tissue Procurement: Ash Alizadeh, John Byrd, M.D., Mike Grever,  
 M.D., Louis M. Staudt, M.D., Ph.D.  
 cDNA Library Preparation: M. Bento Soares, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNT, send email to:  
 info@image.llnl.gov  
 Seg primer: -40UP from Gibco  
 High quality sequence stop: 454.

## FEATURES

source

1.468  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3292830"  
 /issue\_type="B-cell, chronic lymphocytic leukemia"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI CGAP CLL1"  
 /note="Vector: pT73D-Pac (Pharmacia) with a modified  
 polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA  
 was primed with a Not I - oligo(dT) primer [5'  
 TGTTCACATCTGAGAGTGGAGCGGCCCTTCATTTTCTTTT  
 T 3']; double-stranded cDNA was ligated to Eco RI  
 adaptors (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of the modified pT73 vector.  
 Library is normalized, and was constructed by Bento  
 Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 1.4%; Score 44; DB 2; Length 468;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2899 GATCAGCTGAGGCGAGAGTTGAGAGCAGGCTGGCAACATAG 2942  
 DB 241 GATCAGCTGAGGCGAGAGTTGAGAGCAGGCTGGCAACATAG 198

RESULT 138  
 T84449/c 472 bp mRNA linear EST 16-MAR-1995  
 LOCUS yd32d04.r1 Soares fetal liver spleen INFILS Homo sapiens cDNA clone  
 DEFINITION IMAGE:109927 5' similar to contains Alu repetitive element; mRNA  
 sequence.

ACCESSION T84449  
 VERSION T84449  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)

ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homidae; Homo.

REFERENCE 1 (bases 1 to 472)  
 Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawkins, M.,  
 Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Mante, M.,  
 Parsons, J., Rifkin, B., Rohlfing, T., Soares, M., Tan, F.,  
 Trevaeths, E., Waterston, R., Williamson, A., Wohldmann, P. and  
 Wilson, R.  
 The WashU-Merck EST Project  
 Unpublished (1995)

## JOURNAL

## COMMENT

Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@wustl.wustl.edu  
 Insert Size: 862  
 High quality sequence stop: 363 Source: IMAGE Consortium, LLNT  
 This clone is available royalty-free through LLNT; contact the  
 IMAGE Consortium (info@image.llnl.gov) for further information.  
 Insert length: 862 Std Error: 0.00  
 Seg primer: M13RP1  
 High quality sequence stop: 363.

## FEATURES

source

1.472  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:465544"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:109927"  
 /sex="male"  
 /dev\_stage="20 week-post conception fetus"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_lib="Soares fetal liver spleen INFILS"  
 /note="Organ: Liver and Spleen. Vector: pT73D (Pharmacia)  
 with a modified polylinker; Site 1: Pac I; Site 2: Eco RI;  
 1st strand cDNA was primed with a Pac I - oligo(dT) primer  
 [5' AACTGAGAGATTATTAAGATCTTTTCTTTTCTTTT 3']  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Pac I and cloned into the Pac I  
 and Eco RI sites of the modified pT73 vector. Library  
 went through one round of normalization. Library  
 constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 1.4%; Score 44; DB 8; Length 472;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TCCACATGACCTCCAGCTGGGCAAGAGCAAGACTGTCTC 3122  
 DB 138 TCCACATGACCTCCAGCTGGGCAAGAGCAAGACTGTCTC 95

RESULT 139  
 BM983814/c

LOCUS BM983814 473 bp mRNA linear EST 20-FEB-2003  
 DEFINITION UI-CF-DUI-*say-c-24-0-UI*. s1 UI-CF-DUI Homo sapiens cDNA clone  
 UI-CF-DUI-*say-c-24-0-UI* 3', mRNA sequence.  
 ACCESSION BM983814  
 VERSION BM983814.1 GI:19608703  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 473)  
 AUTHORS Donald,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 PUBLISHED 8889548  
 COMMENT Contact: McCray, PB  
 McCray Lab  
 University of Iowa  
 2024 University of Iowa Med Labs, Iowa City, IA 52242, USA  
 Tel: 319 356 4866  
 Fax: 319 356 7171  
 Email: paul.mccray@uiowa.edu  
 Tissue Procurement: Dr. M. J. Welsh, University of Iowa  
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 cDNA Distribution: Researchers may obtain clones from Research  
 Genetics (www.reagen.com) or from Open Biosystems  
 (www.openbiosystems.com).  
 The following repetitive elements were found in this cDNA  
 sequence: 11-219, >ALU (matched complement) 226-461, >LINE2  
 Seq primer: M13 FORWARD  
 POLYA=yes.

FEATURES  
 source  
 Location/Qualifiers  
 1..473  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-CF-DUI-*say-c-24-0-UI*"  
 /tissue\_type="Primary Lung Epithelial Cells"  
 /dev\_stage="Adult"  
 /lab\_host="DH10B (Life Technologies) (T1 phage resistant)"  
 /clone\_1lb="UI-CF-DUI"  
 /note="Organ: Lung; Vector: pTT3-Pac (Pharmacia) with a  
 modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 UI-CF-DUI is a normalized cDNA library containing the  
 following tissue(s): Primary Lung Epithelial Cells The  
 library was constructed according to Donald, Lennon and  
 Soares, Genome Research, 6:791-806, 1996. First strand  
 cDNA synthesis was primed with an oligo-dT primer  
 containing a Not I site. Double stranded cDNA was ligated  
 to an EcoR I adaptor, digested with Not I, and cloned  
 directionally into pTT3-Pac vector. The oligonucleotide  
 used to prime the synthesis of first-strand cDNA contains  
 a library tag sequence that is located between the Not I  
 site and the (dT)18 tail. The sequence tag for this  
 library is GGCTGTAGGC.  
 TAG TISSUE=Lung Epithelial Cells Tissue nos 359-368  
 TAG LIB=UI-CF-DUI  
 TAG\_SEQ=GGCTGTAGGC"

ORIGIN  
 Query Match 1.4%; Score 44; DB 3; Length 473;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 3079 TGGCCTGCACTCCAGCTCGGCAAGAGCAAGACTCTGTCTC 3122  
 |||||||  
 62 TGGCCTGCACTCCAGCTCGGCAAGAGCAAGACTCTGTCTC 19

RESULT 140  
 LOCUS BM314257 483 bp mRNA linear EST 03-JAN-2002  
 DEFINITION 1952c03.x1 HR85 1stet Homo sapiens cDNA 3', mRNA sequence.  
 ACCESSION BM314257  
 VERSION BM314257.1 GI:18048602  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
 Homiidae; Homo.  
 REFERENCE 1 (bases 1 to 483)  
 AUTHORS Melton,D., Brown,J., Kenty,G., Permutt,A., Lee,C., Kaestner,K.,  
 Lemshka,I., Searce,M., Brestelli,J., Gradwohl,G., Clifton,S.,  
 Hillier,L., Maria,M., Page,D., Wylie,T., Martin,J., Blicstein,A.,  
 Schmitt,A., Theising,B., Rutter,E., Ronko,I., Bennett,J.,  
 Cardenas,M., Gibbons,M., McCann,R., Cole,R., Tsagarelshvili,R.,  
 Williams,T., Jackson,Y. and Bowers,Y.  
 Endocrine Pancreas Consortium  
 Unpublished (2000)  
 COMMENT Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue  
 Endocrine Pancreas Consortium  
 Harvard University, Howard Hughes Medical Institute  
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,  
 MA 02138  
 Tel: 617-495-1812  
 Fax: 617-495-8537  
 Email: dmelton@biohp.harvard.edu  
 Library was constructed by Dr. Hiroshi Inoue DNA sequencing by:  
 Washington University Genome Sequencing Center For information on  
 obtaining a clone please contact: Dr. Hiroshi Inoue  
 (hinoue@im.wustl.edu)  
 Seq primer: -400P from Gibco  
 High quality sequence stop: 442.

FEATURES  
 source  
 Location/Qualifiers  
 1..483  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="Purified pancreatic islet"  
 /lab\_host="DH10B"  
 /clone\_1lb="HR85 1stet"  
 /note="Organ: Pancreas; Vector: pBluescript SK(-); Site 1:  
 NotI; Site 2: XhoI; cDNA made by oligo-dT priming.  
 Size-selected on agarose gel. Average insert size -1kb. 5'  
 XhoI site was destroyed after directional cloning.  
 Amplified once. Contact information: Hiroshi Inoue, MD,  
 Metabolism Div. (Alan Permutt Lab), Washington University  
 School of Medicine, Box 8127, 660 South Euclid Ave., St.  
 Louis, MO 63110, E-mail: hinoue@im.wustl.edu, Tel:  
 314-362-1916, Fax: 314-747-2692."

ORIGIN  
 Query Match 1.4%; Score 44; DB 3; Length 483;  
 Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 3073 AGATTGTGCACTGCACTCCAGCTCGGCAAGAGCAAGACTC 3116  
 |||||||  
 377 AGATTGTGCACTGCACTCCAGCTCGGCAAGAGCAAGACTC 420

RESULT 141  
 LOCUS AQ276534 484 bp DNA linear GSS 22-NOV-1998  
 DEFINITION CITBI-B1-2521J15.TR CITBI-B1 Homo sapiens genomic clone 2521J15,  
 genomic survey sequence.  
 ACCESSION AQ276534  
 VERSION AQ276534.1 GI:3902730  
 KEYWORDS GSS.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 484)  
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and Venter,J.C.  
Use of a random human BAC End Sequence Database for Sequence-Ready Map Building  
Unpublished (1998)  
Contact: Mark Adams  
Department of Bukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdamas@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC end search page:  
http://www.tigr.org/cdb/hungen/bac\_end\_search/bac\_end\_search.html.  
Seq primer: M13 Reverse  
Class: BAC ends.  
Location/Qualifiers  
1..484  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="2521015"  
/sex="male"  
/cell\_type="sperm"  
/clone\_1ib="CITBI-81"  
/note="Vector: pBeloBAC11; site\_1: EcoRI; site\_2: EcoRI; Caltech Human BAC Library D"

ORIGIN

Query Match 1.4%; Score 44; DB 9; Length 484;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGACTCTGCTC 3116  
|||||  
Db 267 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGACTCTGCTC 224  
|||||

RESULT 142  
BQ432539 485 bp mRNA linear EST 24-MAY-2002  
LOCUS BQ432539  
DEFINITION AGENCOURT 7905807 NIH\_MGC\_82 Homo sapiens cDNA clone IMAGE:6104737  
5', mRNA sequence.  
ACCESSION BQ432539  
VERSION BQ432539.1 GI:21171615  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 485)  
NIH-MGC http://mgs.nci.nih.gov/.  
National Institutes of Health, Mammalian Gene Collection (MGC)  
Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-rcmail.nih.gov  
Tissue Procurement: CLONTECH  
CDNA Library Preparation: CLONTECH Laboratories, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
http://image.llnl.gov  
Plate: L16M2343 row: b column: 02  
High quality sequence stop: 463.  
Location/Qualifiers

FEATURES

source  
1..485  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:6104737"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_1ib="NIH\_MGC\_82"  
/note="Organ: testis; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggccctcgccg); Site 2: SfiI (ggccctatggcc); 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGCCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCCGAGCGCCGACATG-dt(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.35 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Query Match 1.4%; Score 44; DB 5; Length 485;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TGCCACTGCACCTCCAGCTGGGCAACAGAGACTCTGCTC 3122  
|||||  
Db 371 TGCCACTGCACCTCCAGCTGGGCAACAGAGACTCTGCTC 414  
|||||

RESULT 143  
N35896 486 bp mRNA linear EST 16-JAN-1996  
LOCUS N35896  
DEFINITION YY28C08.s1 Soares melanocyte 2NbhM Homo sapiens cDNA clone IMAGE:272558 3' similar to contains Alu repetitive element, mRNA sequence.  
N35896  
VERSION N35896.1 GI:1157038  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
1 (bases 1 to 486)  
Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J., Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaaskis,E., Waterson,R., Williamson,A., Woldmann,P. and Wilson,R.  
The WashU-Merck EST Project  
Unpublished (1995)  
Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.edu  
High quality sequence stops: 440  
Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.llnl.gov) for further information.  
Seq primer: m13 -40 forward  
High quality sequence stop: 440.  
Location/Qualifiers  
1..486  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:3882200"  
/db\_xref="taxon:9606"  
/clone="IMAGE:272558"  
/sex="Male"  
/tissue\_type="melanocyte"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_1ib="Soares melanocyte 2NbhM"

FEATURES



/note="Vector: pT773D (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer [5', TCTTACCACTCGCACTGCACTCCAGCTGGCAAGAGAGACTC 3'], double-stranded cDNA was size selected, ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of a modified pT773 vector (Pharmacia). Library constructed by Bento Soares and M. Fatima Bonaldo. RNA from normal forebrain melanocytes (FS374) was kindly provided by Dr. Anthony P. Albino."

## ORIGIN

Query Match 1.4%; Score 44; DB 8; Length 486;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACTGCACTCCAGCTGGCAAGAGAGACTC 3116  
DB 380 AGATTGTGCACTGCACTCCAGCTGGCAAGAGAGACTC 423

RESULT 144  
LOCUS A1908381 494 bp mRNA linear EST 30-MAR-2000  
DEFINITION RC-BT170-280399-015 BT170 Homo sapiens cDNA, mRNA sequence.  
ACCESSION A1908381  
VERSION A1908381.1 GI:6499061  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

REFERENCE 1 (bases 1 to 494)  
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M. R., Nagai, M. A., da Silva, W. Jr., Zago, M. A., Bordin, S., Costa, F. F., Goldman, G. H., Carvalho, A. F., Matsukuma, A., Baita, G. S., Simpson, D. H., Brunstein, A., de Oliveira, P. S., Bucher, P., Jongeneel, C. V., O'Hare, M. J., Soares, F., Brentani, R. R., Reis, L. F., de Souza, S. J. and Simpson, A. J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags

TITLE Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

## JOURNAL

## PUBMED

CONTACT: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil

Tel: +55-11-2704932  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/seq/gethtml.pl?lt=RC&ct=RC-BT170-015.html  
ct=3280399&ct=1)  
Seq primer: puc 18 forward.  
FEATURES  
Location/Qualifiers  
1..494

source  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/sex="female"  
/dev\_stage="Adult"  
/clone\_1lb="BT170"  
/note="Organ: breast; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the puc 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

Query Match 1.4%; Score 44; DB 1; Length 494;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACTGCACTCCAGCTGGCAAGAGAGACTC 3116  
DB 378 AGATTGTGCACTGCACTCCAGCTGGCAAGAGAGACTC 421

## RESULT 145

LOCUS R86218/c 495 bp mRNA linear EST 14-AUG-1995  
DEFINITION YP88a10.r1 Soares fetal liver spleen INPFS Homo sapiens cDNA clone IMAGE:194490 5' similar to contains Alu repetitive element;; mRNA sequence.  
ACCESSION R86218  
VERSION R86218.1 GI:944624  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

REFERENCE 1 (bases 1 to 495)  
Hillier, L., Clark, N., Dubuque, T., Elliston, K., Hawking, M., Holman, M., Hultman, M., Kucaba, T., Le, M., Lennon, G., Marra, M., Parsons, J., Rifkin, L., Rohlfing, T., Soares, M., Tan, P., Trevaaskis, E., Waterston, R., Williamson, A., Wohldmann, P. and Wilson, R.  
The WashU-Merck EST Project  
Unpublished (1995)  
CONTACT: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu  
Insert Size: 963  
High quality sequence stops: 379  
Source: IMAGE Consortium, LNL  
This clone is available royalty-free through LNL; contact the IMAGE Consortium (info@image.lnl.gov) for further information.  
Insert length: 963 Std Error: 0.00  
Seq primer: M13P1  
High quality sequence stop: 379.  
Location/Qualifiers  
1..495

## FEATURES

## source

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:3763540"  
/db\_xref="taxon:9606"  
/clone="IMAGE:194490"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_1lb="Soares fetal liver spleen INPFS"  
/note="Organ: liver and spleen; Vector: pT773D (Pharmacia) with a modified polylinker; Site 1: Pac I; Site 2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGAGAGATTAATTAAGATCTTTTCTTTTCTTTTCTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT773 vector. Library went through one round of normalization. Library constructed by Bento Soares and M. Fatima Bonaldo."

## ORIGIN

Query Match 1.4%; Score 44; DB 8; Length 495;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TCCGACATGCACTCCAGCTGGCAAGAGAGAGACTCTGTCTC 3122

Db 255 TCGCAGTGCCTCCAGCTCGGGCAAGCAAGACTCTCTCTC 212

RESULT 146  
LOCUS BE146359 512 bp mRNA linear EST 21-JUN-2000  
DEFINITION BR0-HT0209-010500-110-e11 HT0209 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BE146359  
VERSION BE146359.1 GI:8609083  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

REFERENCE  
AUTHORS 1 (bases 1 to 512)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.U., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

TITLE  
JOURNAL Contact: Simpson A.J.G.  
PUBMED Laboratory of Cancer Genetics  
10737800 Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=et2=MR0-HT0209-010500-110-eltct3-2000-05-01&f4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 126  
High quality sequence stop: 512.  
Location/Qualifiers  
1..512  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="HT0209"  
/note="Organ: head neck; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN  
Query Match 1.4%; Score 44; DB 2; Length 512;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2979 CCGGCGATGATGCGGATGCTGTGTCACGACTCTCGGAGG 3022  
LOCUS BE146359 512 bp mRNA linear EST 08-MAR-2000  
DEFINITION BR0-HT0209-010500-110-e11 HT0209 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BE146359  
VERSION BE146359.1 GI:8609083  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

REFERENCE  
AUTHORS 1 (bases 1 to 512)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.U., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

TITLE  
JOURNAL Contact: Simpson A.J.G.  
PUBMED Laboratory of Cancer Genetics  
10737800 Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=et2=MR0-HT0209-010500-110-eltct3-2000-05-01&f4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 126  
High quality sequence stop: 512.  
Location/Qualifiers  
1..512  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="HT0209"  
/note="Organ: head neck; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

VERSION A1888752.1 GI:5593916  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

REFERENCE  
AUTHORS 1 (bases 1 to 516)  
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D., Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Greg Lennon, Ph.D.  
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/ILNI at: www.bio.inl.gov/bbrp/image/image.html  
Insert Length: 1881 Sd Error: 0.00  
Seq primer: ~40UP from G1bco  
High quality sequence stop: 413.  
Location/Qualifiers  
1..516  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="IMAGE:2447508"  
/tissue\_type="poorly differentiated adenocarcinoma with signet ring cell features"  
/lab\_host="DH10B"  
/clone\_lib="NCI-CGAP\_Ga4"  
/note="Organ: stomach; Vector: pCMV-SPORT6; Site 1: SalI; Site 2: NciI; Cloned unidirectionally. Primer: Oligo dt. Average insert size 1.69 kb. Life Technologies catalog #: 11549-011"

ORIGIN  
Query Match 1.4%; Score 44; DB 1; Length 516;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 3073 AATATTGCGCACTGCCTCCAGCTCGGGCAAGCAAGACTC 3116  
LOCUS BE940436 517 bp mRNA linear EST 02-OCT-2000  
DEFINITION RC3-UT0062-210800-021-c05 UT0062 Homo sapiens cDNA, mRNA sequence.  
ACCESSION BE940436  
VERSION BE940436.1 GI:10469931  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.

REFERENCE  
AUTHORS 1 (bases 1 to 517)  
Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Briones,M.R., Nagai,M.A., da Silva,M. Jr., Zago,M.A., Bordin,S., Costa,F.F., Goldman,G.H., Carvalho,A.F., Matsukuma,A., Bala,G.S., Simpson,D.H., Brunstein,A., deoliveira,P.S., Bucher,P., Jongeneel,C.V., O'Hare,M.U., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and Simpson,A.J.  
Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

TITLE  
JOURNAL Contact: Simpson A.J.G.  
PUBMED Laboratory of Cancer Genetics  
10737800 Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=et2=MR0-HT0209-010500-110-eltct3-2000-05-01&f4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 126  
High quality sequence stop: 512.  
Location/Qualifiers  
1..512  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="HT0209"  
/note="Organ: head neck; Vector: puc18; Site 1: SmaI; Site 2: SmaI; A mini-library was made by cloning products derived from ORESTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## COMMENT

Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001  
Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome  
Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?l=kt2=RC3-UT0062-210  
800-021-c05kt3=2000-08-21kt4=1)  
Seq primer: puc 18 forward  
High quality sequence start: 9  
High quality sequence stop: 517.  
Location/Qualifiers

## FEATURES

## source

1..517  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="adult"  
/clone\_lib="UT0062"  
/note="Organ: uterus tumor; Vector: puc18; Site\_1: SmaI;  
Site\_2: SmaI; A mini-library was made by cloning products  
derived from ORESTES PCR (U.S. Letters Patent application  
No. 196,716 - Ludwig Institute for Cancer Research)  
profiles into the pUC 18 vector. Reverse transcription of  
tissue mRNA and cDNA amplification were performed under  
low stringency conditions."

## ORIGIN

## Query Match

Best Local Similarity 1.4%; Score 44; DB 2; Length 517;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TGGCAGTGCATCTCAGCTGGGCAAGCAAGCAAGCTCTCTC 3122  
DB 354 TGGCAGTGCATCTCAGCTGGGCAAGCAAGCAAGCTCTCTC 311

## RESULT 149

## LOCUS

AQ026637

## DEFINITION

527 bp DNA linear GSS 30-JUN-1998  
CIT-HSP-2314017.TR CIT-HSP Homo sapiens genomic clone 2314017,  
genomic survey sequence.

## ACCESSION

## VERSION

AQ026637.1 GI:3266859

## KEYWORDS

GSS.

## SOURCE

Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

## REFERENCE

## AUTHORS

1 (bases 1 to 527)  
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bases,S., Linher,K.,  
Golden,K., Berry,K., Granger,D., Suh,B., Wible,C., Shizuya,H.,  
Simon,M. and Venter,A.C.  
Use of a random BAC End Sequence Database for Sequence-Ready Map  
Building (1998)

## TITLE

## JOURNAL

## COMMENT

Unpublished (1998)  
Other GSSs: CIT-HSP-2314017.TP  
Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdadam@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC  
end search page:  
http://www.tigr.org/tdb/humgen/bac\_end\_search/bac\_end\_search.html.  
Seq primer: M13 Reverse  
Clase: BAC ends.

## FEATURES

## source

Location/Qualifiers  
1..527  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone\_lib="2314017"  
/sex="Male"  
/cell\_type="Sperm"  
/clone\_lib="CIT-HSP"  
/note="Vector: pBeloBAC11; Site\_1: HindIII; Site\_2:  
HindIII"

## ORIGIN

## Query Match

Best Local Similarity 1.4%; Score 44; DB 9; Length 527;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGAGGTGATCACTGAGGCCAGCAAGTTCGAGCAAGCCTG 2932  
DB 215 GAGGAGGTGATCACTGAGGCCAGCAAGTTCGAGCAAGCCTG 258

## RESULT 150

## LOCUS

CN389220

## DEFINITION

530 bp mRNA linear EST 16-MAY-2004  
1700060037068 GRN\_PRENBU Homo sapiens cDNA 5', mRNA sequence.

## ACCESSION

CN389220

## VERSION

CN389220.1 GI:47376815

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

## ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

## REFERENCE

## AUTHORS

1 (bases 1 to 530)  
Brandenberger,R., Wei,H., Zhang,S., Lei,S., Murage,J., Flak,G.J.,  
Li,Y., Xu,C., Pang,R., Guegler,K., Rao,M.S., Mandalam,R.,  
Lebkowski,J and Stanton,L.W.  
Transcriptome characterization elucidates signaling networks that  
control human ES cell growth and differentiation  
Nat. Biotechnol. 22 (6), 707-716 (2004)

## JOURNAL

## PUBMED

15146197

## COMMENT

Contact: Brandenberger R

Regenerative Medicine  
Geron Corporation  
230 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8638  
Fax: 650 473 7760  
Email: rbrandenberger@geron.com

## FEATURES

Insert Length: 530 Std Error: 0.00.

## source

Location/Qualifiers

1..530  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/ligase\_type="embryonic stem cell, retinoic acid and  
mitogen-treated hES cell line H7"  
/clone\_lib="GRN\_PRENBU"  
/note="Oligo dT primed, full-length enriched cDNA library  
from hES cell line H7 (p29) maintained in feeder-free  
conditions. Embryoid bodies were generated in the presence  
of all-trans retinoic and mitogens."

## ORIGIN

## Query Match

Best Local Similarity 1.4%; Score 44; DB 7; Length 530;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCTGCTGCTCAGCTGGGCAAGCAAGCAAGCTC 3116  
DB 416 AGATTGTGCTGCTGCTCAGCTGGGCAAGCAAGCAAGCTC 373

RESULT 151  
LOCUS CD691041 538 bp mRNA linear EST 25-JUN-2003  
DEFINITION EST5564 human nasopharynx Homo sapiens cDNA, mRNA sequence.  
ACCESSION CD691041  
VERSION CD691041.1 GI:32212370  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 538)  
AUTHORS Liu X.-Q., Zhou Y., Zhang L.-J., Xu H., Chen H.-K., Pan Z.-G. and Zeng Y.-X.  
TITLE Transcriptional Gene Expression Profile of Human Nasopharynx  
JOURNAL Unpublished (2003)  
COMMENT Contact: Yixin Zeng  
Cancer Center  
Sun Yat-sen University  
651 Dongfeng Road East, Guangzhou 510060, China  
Tel: 86-1380-9770-743  
Fax: 86-20-8775-4506  
Email: yxzeng@zsums.edu.cn.  
Location/Qualifiers  
1..538  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue\_type="normal nasopharynx"  
/clone\_lib="human nasopharynx"  
/note="ESTs generated from a normal nasopharynx cDNA library from southern Chinese"

ORIGIN  
Query Match 1.4%; Score 44; DB 6; Length 538;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TGCCACTGCACCTCCAGCTGGGCAAGAGCAAGACTCTGTCTC 3122  
|||||  
|||||  
Db 326 TGCCACTGCACCTCCAGCTGGGCAAGAGCAAGACTCTGTCTC 369

RESULT 152  
LOCUS CV418654/C 540 bp mRNA linear EST 28-SEP-2004  
DEFINITION RC3-UT0062-210800-021-b11 UT0062 Homo sapiens cDNA, mRNA sequence.  
ACCESSION CV418654  
VERSION CV418654.1 GI:52814157  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 540)  
AUTHORS Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R., Nagai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Coêta, F.F., Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bal, G.S., Simpson, D.H., Bruneteau, A., deOliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and Simpson, A.J.  
TITLE Shotgun sequencing of the human transcriptome with ORF expressed sequence tags  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
PUBMED 10737800  
COMMENT Contact: Simpson A.J.G.  
Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP, Brazil  
Tel: +55-11-2704922

FEATURES  
source 1..540  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_lib="UT0062"  
/note="Organ: uterus tumor; Vector: puc18; site 1: SmaI; site 2: SmaI; A mini-library was made by cloning products derived from ORG875 PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

ORIGIN  
Query Match 1.4%; Score 44; DB 7; Length 540;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TGCCACTGCACCTCCAGCTGGGCAAGAGCAAGACTCTGTCTC 3122  
|||||  
|||||  
Db 367 TGCCACTGCACCTCCAGCTGGGCAAGAGCAAGACTCTGTCTC 324

RESULT 153  
LOCUS B55060/C 543 bp DNA linear GSS 20-JUN-1998  
DEFINITION CIT-HSP-385P14.TF CIT-HSP Homo sapiens genomic clone 385P14,  
B55060 genomic survey sequence.  
ACCESSION B55060  
VERSION B55060.1 GI:2609394  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homiidae; Homo.  
REFERENCE 1 (bases 1 to 543)  
AUTHORS Adams, M.D., Rounsley, S.D., Field, C.E., Bass, S., Linher, K., Golden, K., Berry, K., Granger, D., Suh, B., Wible, C., Shizuya, H., Simon, M., and Venter, J.C.  
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map Building  
JOURNAL Unpublished (1997)  
COMMENT Other GSSs: CIT-HSP-385P14.TR  
Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdamas@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC end search page:  
http://www.tigr.org/cdb/hungen/bac\_end\_search/bac\_end\_search.html  
Seq primer: M13-21  
Class: BAC ends.  
Location/Qualifiers  
1..543  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:5378919"  
/db\_xref="taxon:9606"  
/clone\_lib="385P14"  
/sex="Male"  
/cell\_type="Sperm"  
/clone\_lib="CIT-HSP"  
/note="Vector: pBelobAC11; site\_1: HindIII; site\_2:

ORIGIN HindIII"

Query Match 1.4%; Score 44; DB 9; Length 543;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGCGACCTGCGCTGGGCAAGACGACATC 3116  
|||||  
DB 436 AGATTGCGACCTGCGCTGGGCAAGACGACATC 393

RESULT 154  
AA715817/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
COMMENT

AA715817 543 bp DNA linear GSS 13-JUL-1999  
HS\_5399\_B1.D03 SP68 RPCT-11 Human Male BAC Library Homo sapiens  
genomic clone Plate=975 Col=5 Row=H, genomic survey sequence.  
AA715817.1 GI:5462987  
GSS;  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
1 (bases 1 to 543)  
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,  
Keller,A., Shaker,R., Furlong,T., Young,T., Zhao,S., Adams,W.D. and  
Hood,L.  
Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
10449764  
Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones are derived from the human BAC library RPCT-11. For BAC  
library availability, please contact Plier de Jong  
(plierdejong.med.buffalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering\_bac.html)  
or from Research Genetics (info@resgen.com). BAC end Web server:  
http://www.htec.washington.edu  
Plate: 975 row: H column: 5  
Seq primer: SP6  
Clase: BAC ends  
High quality sequence stop: 543.

# FEATURES

source  
1. 543  
Location/Qualifiers

/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="Plate=975 Col=5 Row=H"  
/sex="male"  
/clone\_1lb="RPCT-11 Human Male BAC Library"  
/note="Vector: pBACe3.6; Site 1: EcoRI; Site 2: EcoRI;  
Male blood DNA was isolated from one randomly chosen donor  
and partially digested with a combination of EcoRI and  
EcoRI Methyase. Size selected DNA was cloned into the  
pBACe3.6 vector at EcoRI sites"

## ORIGIN

Query Match 1.4%; Score 44; DB 9; Length 543;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TGGCACTGCACTCAGCTGGGCAAGACGACATCTCTCTC 3122  
|||||  
DB 391 TGGCACTGCACTCAGCTGGGCAAGACGACATCTCTCTC 348

RESULT 155  
AA715817/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
COMMENT

AA715817 551 bp mRNA linear EST 22-JAN-1998  
nm23c05.s1 NCI CGAP GCBO Homo sapiens CDNA clone IMAGE:1241480 3  
similar to gb:K01144 HLA CLASS II HISTOCOMPATIBILITY ANTIGEN, GAMMA  
CHAIN PRECURSOR (HUMAN); contains Alu repetitive element; contains  
element PRT5 repetitive element; , mRNA sequence.  
AA715817.1 GI:2728091  
EST.  
Homo sapiens  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
1 (bases 1 to 551)  
NCI CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cga@bbs-rcmail.nih.gov  
Tissue Procurement: Louis M. Staudt, M.D., Ph.D., David Allman,  
Ph.D., Gerald Marti, M.D.  
CDNA Library Preparation: Louis M. Staudt, M.D., Ph.D.  
DNA Sequencing by: Washington University Genome Sequencing Center  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/ILN at:  
www.bio.lnlnl.gov/bbrp/image/image.html  
Insert length: 1083 Std Error: 0.00  
Seq primer: -40m3 fwd. RT from Amersham  
High quality sequence stop: 352.

# FEATURES

source  
1. 551  
Location/Qualifiers

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:1241480"  
/tissue\_type="germinal center B-cells"  
/lab\_host="SOLR (kanamycin resistant)"  
/clone\_1lb="NCI CGAP GCBO"  
/note="Organ: tonsil; Vector: Bluescript SK-; Site 1:  
EcoRI; Site 2: XhoI; Cloned unidirectionally. Primer:  
Oligo dr. Germinal center B-cells library constructed by  
Dr. L. Staudt (NCI). 5' adaptor sequence: 5'  
GAATTCGCGACGAG 3' 3' adaptor sequence: 5'  
CTCGAGTTTCTTTTCTTTT 3' Average insert size: 1.1 kb."

## ORIGIN

Query Match 1.4%; Score 44; DB 1; Length 551;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2899 GATCAGCTGAGGCGAGGTTGAGACGAGCTGGGCAACATAG 2942  
|||||  
DB 248 GATCAGCTGAGGCGAGGTTGAGACGAGCTGGGCAACATAG 205

RESULT 156  
CB129216/c  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
PUBMED  
COMMENT

CB129216 558 bp mRNA linear EST 29-JAN-2003  
K-EST0178834 L15CKX1 Homo sapiens CDNA clone L15CKX1-1-B09 5', mRNA  
sequence.  
CB129216  
CB129216.1 GI:28092697  
EST.  
Homo sapiens (human)  
Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.

Laboratory of Cancer Genetics  
Ludwig Institute for Cancer Research  
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP  
Brazil  
Tel: +55-11-2704922  
Fax: +55-11-2707001

Email: asimpson@ludwig.org.br  
This sequence was derived from the FAPESP/LICR Human Cancer Genome Project. This entry can be seen in the following URL  
(http://www.ludwig.org.br/scripts/gethtml2.pl?cl=8&c2=MR3-OT0005-170  
200-101-c01&c3=2000-02-17&c4=1)  
Seq primer: puc 18 forward  
High quality sequence stop: 337.

## FEATURES

source

1.563  
Location/Qualifiers

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/dev\_stage="Adult"  
/clone\_1ib="OT0005"  
/note="Organ: ovary; Vector: puc18; Site\_1: SmaI; Site\_2: SmaI; A mini-library was made by cloning products derived from ORSTES PCR (U.S. Letters Patent application No. 196,716 - Ludwig Institute for Cancer Research) profiles into the pUC 18 vector. Reverse transcription of tissue mRNA and cDNA amplification were performed under low stringency conditions."

## ORIGIN

Query Match 1.4%; Score 44; DB 1; Length 563;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TCCACCTGCATCCAGCTGGGCAACAGACGACTGTCTC 3122  
|||||  
DB 263 TCCACCTGCATCCAGCTGGGCAACAGACGACTGTCTC 306

## RESULT 159

LOCUS A0591303 569 bp DNA linear GSS 08-JUN-1999  
DEFINITION HS 2122 B1 G04 MR CIT Approved Human Genomic Sperm Library D Homo  
sapiens genomic clone Plate=2122 Col=7 Row=N, genomic survey  
sequence.

ACCESSION A0591303  
VERSION A0591303.1 GI:5022955  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 569)  
AUTHORS Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,  
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and  
Hood,L.  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

TITLE Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
PUBMED 10449764

COMMENT Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887

Email: jwallace@u.washington.edu  
Clones may be purchased from Research Genetics (info@resgen.com).  
BAC end Web Server: http://www.htsc.washington.edu  
Plate: 2122 Row: N Column: 7  
Seq primer: M13 Reverse  
Clas: BAC ends  
High quality sequence stop: 569.

## FEATURES

source

1.569  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="Plate=2122 Col=7 Row=N"

/sex="male"  
/clone\_1ib="CIT Approved Human Genomic Sperm Library D"  
/note="Organ: sperm; Vector: pBelOBAC11; BAC Clones in  
E-Coli DH10B"

## ORIGIN

Query Match 1.4%; Score 44; DB 9; Length 569;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCGAGTGATCCTGAGCCGAGAGTTGAGACGACCTG 2932  
|||||  
DB 266 GAGGCGAGTGATCCTGAGCCGAGAGTTGAGACGACCTG 309

RESULT 160  
LOCUS CD723228/C 581 bp mRNA linear EST 26-JUN-2003  
DEFINITION o19e04.y1 Human lacrimal gland, unamplified: o19e04 5', mRNA sequence.  
CD723228 clone o19e04 5', mRNA sequence.

ACCESSION CD723228.1 GI:32274076  
VERSION CD723228.1  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 581)  
AUTHORS Ozyildirim,A.M., Wistow,G.J., Gao,J.J., Wang,J., Dickinson,D.P.,  
Frieson,H.F. Jr and Laurie,G.W.  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.

TITLE The lacrimal gland transcriptome is an unusually rich source of  
rare and poorly characterized gene transcripts  
JOURNAL Invest. Ophthalmol. Vis. Sci. 46 (5), 1572-1580 (2005)  
PUBMED 15851553

COMMENT Contact: Wistow G  
Section on Molecular Structure and Function  
National Eye Institute  
6/331, NIH, Bethesda, MD 20892-2740, USA  
Tel: 301 402 3452  
Fax: 301 496 0078  
Email: gwaene@helix.nih.gov  
Plate: 19 row: e column: 04  
Seq primer: M13R1 reverse primer (ABI).  
Location/Qualifiers

## FEATURES

source

1.581  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="O19e04"  
/tissue\_type="Lacrimal gland"  
/dev\_stage="Adult"  
/lab\_host="EMDH10B"  
/note="Organ: Eye; Vector: pCMVSPORT6; RNA was extracted  
from 2 human lacrimal glands. A directionally cloned cDNA  
library in the pCMVSPORT6 vector (Life Technologies) was  
constructed at Bioserve Biotechnology (Laurel MD)  
essentially following the protocols of the Superscript  
Plasmid System full details of which are contained in the  
manufacturer's instruction manual  
(http://www.lifetech.com/). First strand synthesis was  
carried out using a Not I primer-adaptor  
[5'-pGACTAGTTCATGATGCGAGCGGCGCCCT(T)15-3']. EST analysis  
was performed on the unamplified library at the NIH  
Intramural Sequencing Center (NISC)."

## ORIGIN

Query Match 1.4%; Score 44; DB 6; Length 581;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGCGCACTCCAGCTGGGCAACAGACGACTC 3116



Db 138 AGATTGTGCACTGCATCTCCAGCTGGGCAACAGACGAACTC 95

RESULT 161  
LOCUS CB218356/c  
DEFINITION NISC nb08c01.x1 COGENE 6E MAN Homo sapiens cDNA clone IMAGE:5795329  
3', mRNA sequence.  
ACCESSION CB218356  
VERSION CB218356.1 GI:28266548  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 582)  
AUTHORS Homidiae, Homo.  
TITLE NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.  
JOURNAL National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
COMMENT Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
CDNA Library Preparation:  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LINL  
DNA Sequencing by: National Institutes of Health Intramural  
Sequencing Center (NISC)  
Clone distribution: NCI-CCAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LINL at:  
info@image.llnl.gov  
Place: LLNL12898 row: F column: 2  
Seq primer: -21M3 forward primer (ABT).  
Location/Qualifiers

FEATURES  
source  
1..582  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5795329"  
/tissue\_type="mandible, pooled"  
/dev\_stage="embryo, 6 weeks postconception"  
/lab\_host="DH10B"  
/clone\_1ib="COGENE 6E MAN"  
/note="Vector: PAMPI; cDNA primed using oligo-dT primer,  
directionally cloned into UDG sites of PAMPI. Size  
selected for insert sizes ranging from 0.2-2.0 kb.  
Normalized to Cots. Primary library, non-amplified.  
Library constructed by M. Lovett. For more information on  
this library, please contact R. Tidwell (Washington  
University) or visit the COGENE website at  
http://hg.wustl.edu/COGENE/."

## ORIGIN

Query Match 1.4%; Score 44; DB 6; Length 582;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACTGCATCTCCAGCTGGGCAACAGACGAACTC 3116  
|||||  
DB 93 AGATTGTGCACTGCATCTCCAGCTGGGCAACAGACGAACTC 50

RESULT 162  
LOCUS AA206418/c  
DEFINITION AA206418 590 bp mRNA linear EST 12-MAR-1998  
clone IMAGE:645104 3' similar to contains Alu repetitive  
element; contains element MER6 repetitive element; mRNA sequence.  
ACCESSION AA206418 GI:1801856  
VERSION EST.  
KEYWORDS Homo sapiens (human)  
SOURCE Homo sapiens

REFERENCE 1 (bases 1 to 590)  
AUTHORS Hillier, L., Allen, M., Bowles, L., Dubuque, T., Geisel, G., Joet, S.,  
Kizman, D., Kucaba, T., Lacy, M., Le, N., Lennon, G., Maier, M.,  
Martin, J., Moore, B., Schellenberg, K., Stepec, M., Tan, F.,  
Theising, B., White, Y., Wylie, T., Waterston, R. and Wilson, R.

TITLE WashU-NCI human EST project  
JOURNAL Unpublished (1997)  
COMMENT Contact: Wilson RK  
Washington University School of Medicine  
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
Tel: 314 286 1800  
Fax: 314 286 1810  
Email: est@wustl.wustl.edu

FEATURES  
source  
1..590  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="GDB:521535"  
/db\_xref="taxon:9606"  
/clone="IMAGE:645104"  
/dev\_stage="Ntera-2/RA neuroepithelial cells"  
/lab\_host="SOLR (kanamycin resistant)"  
/clone\_1ib="Stratagene neuroepithelium (#937231)"  
/note="Vector: pBluescript SK-; Site 1: EcoRI; Site 2:  
XhoI; Cloned unidirectionally. Primer: Oligo dT. NT2  
cells (Ntera-2/cl.D1) induced with Retinoic Acid for 24  
hours. Average insert size: 1.5 kb; Uni-ZAP XR Vector; -5'  
adaptor sequence: 5' GAATTGCGACGAG 3' -3' adaptor  
sequence: 5' CTCGAGTTTCTTTTCTT 3'."

## ORIGIN

Query Match 1.4%; Score 44; DB 1; Length 590;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACTGCATCTCCAGCTGGGCAACAGACGAACTC 3116  
|||||  
DB 68 AGATTGTGCACTGCATCTCCAGCTGGGCAACAGACGAACTC 25

RESULT 163  
LOCUS B70450  
DEFINITION CIT-HSP-2059C13.TF CIT-HSP Homo sapiens genomic clone 2059C13,  
genomic survey sequence.  
ACCESSION B70450  
VERSION B70450.1 GI:2709674  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 602)  
AUTHORS Adams, M.D., Rounseley, S.D., Field, C.E., Bass, S., Linher, K.,  
Simon, M., and Venter, J.C.  
TITLE Use of a random BAC End Sequence Database for Sequence-Ready Map  
Building  
JOURNAL Unpublished (1997)  
COMMENT Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200

Fax: 301 838 0208  
Email: mdadams@tigr.org  
Clones are available from Research Genetics (info@resgen.com). BAC  
end search page:  
http://www.tigr.org/cdb/humgen/bac\_end\_search/bac\_end\_search.html  
Seq primer: M13-21  
Class: BAC ends.

# FEATURES

source

1..602  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/db\_xref="taxon:9606"  
/clone="2059C13"  
/sex="Male"  
/cell\_type="Sperm"  
/clone\_1fb="CIT-HSP"  
/note="Vector: pBelBAC11, Site\_1: HindIII, Site\_2:  
HindIII"

## ORIGIN

Query Match 1.4%; Score 44; DB 9; Length 602;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACATGCACTCCAGCCTGGGCAACAGACCAAGACTC 3116  
|||||  
DB 11 AGATTGTGCACATGCACTCCAGCCTGGGCAACAGACCAAGACTC 54

## RESULT 164

AL691931 614 bp mRNA linear EST 04-SEP-2003  
LOCUS DKF2P313B1436.r1 313 (synonym: h1ccc2) Homo sapiens cDNA clone  
DEFINITION DKF2P313B1436 5', mRNA sequence.  
ACCESSION AL691931  
VERSION AL691931.1 GI:19617508  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 614)  
Koehler,K., Beyer,A., Mewes,W., Weil,B. and Wiemann,S.  
EST (Koehler,K., Beyer,A., Mewes,W., Weil,B. and Wiemann,S.)  
Unpublished (1999)  
COMMENT Contact: MIPS  
MIPS

Ingolstedter Landstr.1, D-85764 Neuherberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;  
sequenced by BMFZ (Biomedical Research Center at the Charite,  
Berlin/Germany) within the cDNA sequencing consortium of the German  
Genome Project.  
No 5' sequence available.  
This clone (DKF2P313B1436) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
Location/Qualifiers  
1..614  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKF2P313B1436"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_1fb="313 (synonym: h1ccc2)"  
/note="Vector: pTriplex2, Site\_1: SfiI, Site\_2: SfiIB,  
cDNA-collection"

## FEATURES

source

1..614  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKF2P313B1436"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_1fb="313 (synonym: h1ccc2)"  
/note="Vector: pTriplex2, Site\_1: SfiI, Site\_2: SfiIB,  
cDNA-collection"

## ORIGIN

Query Match 1.4%; Score 44; DB 1; Length 614;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TGCACATGCACTCCAGCCTGGGCAACAGACCAAGACTCTGTCTC 3122  
|||||  
DB 304 TGCACATGCACTCCAGCCTGGGCAACAGACCAAGACTCTGTCTC 347

## RESULT 165

CL246302 614 bp DNA linear GSS 22-JAN-2004  
LOCUS HSC 01025 RPCT-11 Human Male BAC Library Homo sapiens genomic clone  
DEFINITION RP11-441N19, genomic survey sequence.  
ACCESSION CL246302.1 GI:41102792  
VERSION CL246302.1 GI:41102792  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 614)  
Scherer,S.W., Rommens,J.M. and Tsui,L.C.  
Gene identification on Human Chromosome 7 (Jan 05/04)  
Unpublished (2003)  
Contact: Scherer, S.W., Rommens, J.M.; Tsui, L.C.  
The Human Chromosome 7 Project  
Department of Genetics, The Hospital for Sick Children  
555 University Avenue, Toronto, Ontario M5G 1X8, Canada  
Tel: 416 813 7613  
Fax: 416 813 8319  
Email: steve@genet.sickkids.on.ca  
Maps to human chromosome 7q11.23  
Seq primer: T7  
Classes: BAC ends  
High quality sequence stop: 614.  
Location/Qualifiers  
1..614  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="RP11-441N19"  
/sex="male"  
/clone\_1fb="RPCT-11 Human Male BAC Library"  
/note="Vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
Male blood DNA was isolated from one randomly chosen donor  
and partially digested with a combination of EcoRI and  
EcoRI Methylase. Size selected DNA was cloned into the  
pBAC3.6 vector at EcoRI sites"

REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
COMMENT

## FEATURES

source

1..614  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="RP11-441N19"  
/sex="male"  
/clone\_1fb="RPCT-11 Human Male BAC Library"  
/note="Vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
Male blood DNA was isolated from one randomly chosen donor  
and partially digested with a combination of EcoRI and  
EcoRI Methylase. Size selected DNA was cloned into the  
pBAC3.6 vector at EcoRI sites"

## ORIGIN

Query Match 1.4%; Score 44; DB 10; Length 614;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3057 GAGGTTGCAATGAGCAAGATTGTGCACATGCACTCCAGCTGGG 3100  
|||||  
DB 546 GAGGTTGCAATGAGCAAGATTGTGCACATGCACTCCAGCTGGG 589

## RESULT 166

AL120523/c 621 bp mRNA linear EST 04-SEP-2003  
LOCUS DKF2P761004.8.g1 761 (synonym: hamy2) Homo sapiens cDNA clone  
DEFINITION DKF2P761004.8 3', mRNA sequence.  
ACCESSION AL120523  
VERSION AL120523.1 GI:59264422  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

## FEATURES

source

1..621  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKF2P761004.8"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_1fb="761 (synonym: hamy2)"  
/note="Vector: pTriplex2, Site\_1: SfiI, Site\_2: SfiIB,  
cDNA-collection"

## ORIGIN

REFERENCE  
1 (bases 1 to 621)  
AUTHORS  
Koehler, K., Beyer, A., Mewes, H.W., Gassenhuber, J. and Wiemann, S.  
TITLE  
EST (Koehler, et al.)  
JOURNAL  
Unpublished (1999)  
COMMENT  
Contact: MIPS  
MIPS

Ingolstaedter Landstr.1, D-85764 Neuberg, Germany  
This is the 3' sequence of the clone insert.  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;  
sequenced by BMFZ (Biomedical Research Center at the Charite,  
Berlin/Germany) within the cDNA sequencing consortium of the German  
Genome Project.  
ri sequence also available.  
This clone (DKFZp7610048) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES  
source  
1..621  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZp7610048"  
/issue\_type="amygdala"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_1lb="761 (synonym: hamy2)"  
/note="Vector: pSPORT1; Site\_1: NotI; Site\_2: SalI"

## ORIGIN

Query Match 1.4%; Score 44; DB 1; Length 621;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
3071 CAAGATTGCGCACTGCAGCTCGGCGCAAGAGCAAGAC 3114  
151 CAAGATTGCGCACTGCAGCTCGGCGCAAGAGCAAGAC 108

RESULT 167  
LOCUS  
CN267712 627 bp mRNA linear EST 16-MAY-2004  
DEFINITION  
17000531627977 GRN\_EB Homo sapiens cDNA 5', mRNA sequence.  
ACCESSION  
CN267712  
VERSION  
CN267712.1 GI:47284126  
KEYWORDS  
EST.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 627)

REFERENCE  
AUTHORS  
Brandenberger, R., Wei, H., Zhang, S., Lei, S., Muraige, J., Flak, G.J.,  
Li, Y., Xu, C., Pang, R., Guegler, K., Rao, M.S., Mandalam, R.,  
Lebowicki, J. and Stanton, L.W.  
TITLE  
Transcriptome characterization elucidates signaling networks that  
control human ES cell growth and differentiation  
JOURNAL  
Nat. Biotechnol. 22 (6), 707-716 (2004)  
PUBMED  
15146197  
COMMENT  
Contact: Brandenberger R  
Regenerative Medicine  
330 Constitution Drive, Menlo Park, CA 94025, USA  
Tel: 650 473 8658  
Fax: 650 473 7760  
Email: rbrandenberger@geron.com

FEATURES  
source  
1..627  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
Location/Qualifiers  
1..627  
Insert Length: 627 Std Error: 0.00.

/tissue\_type="embryonic stem cells, embryoid bodies  
derived from H1, H7 and H9 cells"  
/clone\_1lb="GRN\_EB"  
/note="Oligo dt primed, full-length enriched cDNA library  
from embryoid body outgrowths derived from h9s cell lines  
H1 (p32), H7 (p29), and H9 (p26) maintained in feeder-free  
conditions."

## ORIGIN

Query Match 1.4%; Score 44; DB 7; Length 627;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
3073 AGATTGTGCACCTGCAGCTCGGCGCAAGAGCAAGACTC 3116  
583 AGATTGTGCACCTGCAGCTCGGCGCAAGAGCAAGACTC 626

RESULT 168  
LOCUS  
A0542374/c 633 bp DNA linear GSS 19-MAY-1999  
DEFINITION  
RPCI-11-346112.TV RPCI-11 Homo sapiens genomic clone  
RPCI-11-346112, genomic survey sequence.  
ACCESSION  
A0542374  
VERSION  
A0542374.1 GI:4872830  
KEYWORDS  
GSS.  
SOURCE  
Homo sapiens (human)  
ORGANISM  
Homo sapiens  
Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 633)  
Zhao, S., Adams, M.D., Nierman, W., Malek, J., de Jong, P. and  
Venter, J.C.  
TITLE  
Use of BAC End Sequences from Library RPCI-11 for Sequence-Ready  
Map Building  
JOURNAL  
Unpublished (1997)  
COMMENT  
Other GSS: RPCI-11-346112.TV  
Contact: Shaying Zhao, William Nierman, Mark Adams  
Department of Bukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: hbe@tigr.org  
Clones are derived from the human BAC library RPCI-11. For BAC  
library availability, please contact Pieter de Jong  
(pieter@dejong.med.buifalo.edu). Clones may be purchased from  
BACPAC Resources (http://bacpac.med.buifalo.edu/ordering) or from  
Research Genet. cs (info@resgen.com). BAC end search page:  
http://www.tigr.org/cdb/hungen/bac\_end\_search/bac\_end\_search.html.  
Seq primer: 586  
Classes: BAC ends.

FEATURES  
source  
1..633  
Location/Qualifiers  
1..633  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="GDB:7632683"  
/db\_xref="taxon:9606"  
/clone="RPCI-11-346112"  
/sex="Male"  
/cell\_type="Lymphocytes"  
/clone\_1lb="RPCI-11"  
/note="Vector: pBACe3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
RPC11 Human Male BAC Library"

## ORIGIN

Query Match 1.4%; Score 44; DB 9; Length 633;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
3078 GTGCCACTGCAGCTCGGCGCAAGAGCAAGACTGTGCT 3121  
|||||

DB 507 GNGCAGTGCAGCTCCAGCTGGGCAACAGAGCAAGACTGTCT 464

RESULT 169

LOCUS BX482192 637 bp mRNA linear EST 04-SEP-2003

DEFINITION DKFZp666H14229.r1.686 (synonym: h1cc3) Homo sapiens cDNA clone

ACCESSION BX482192

VERSION BX482192.1 GI:31941851

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 637)

AUTHORS Bloecker, H., Boecker, M., Mewes, H.W., Well, B., Amlid, C., Oeanger, A., Pobo, G., Han, M., and Wiemann, S.

TITLE EST (Bloecker, H., Boecker, M., Mewes, H.W., Well, B., Amlid, C., et al.)

JOURNAL Unpublished (2003)

COMMENT Contact: MIPS

MIPS Ingolstaedter Landstr.1, D-85764 Neuberg, Germany

This is the 5' sequence of the clone insert

Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de; sequenced by GPF (National Research Centre for Biotechnology Ltd., Braunschweig/Germany) within the cDNA sequencing consortium of the German Genome Project.

No al sequence available.

This clone (DKFZp666H14229) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

Location/Qualifiers

1..637

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="DKFZp666H14229"

/dev\_stage="adult"

/lab\_host="DH10B"

/clone\_1lb="686 (synonym: h1cc3)"

/note="Vector: pT7.1pLX2; Site\_1: SfiI; Site\_2: SfiI; cDNA-collection"

ORIGIN

Query Match 1.4%; Score 44; DB 5; Length 637;

Best Local Similarity 100.0%; Pred. No. 1.9e-09;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TGCCAGTGCAGCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122

DB 389 TGCCAGTGCAGCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 432

RESULT 170

LOCUS AQ697816 638 bp DNA linear GSS 06-JUL-1999

DEFINITION HS 5536.B1.B10.T7A.RPCI-11 Human Male BAC library Homo sapiens genomic clone Plate=1112 Col=19 Row=D, genomic survey sequence.

ACCESSION AQ697816

VERSION AQ697816.1 GI:5388064

KEYWORDS GSS.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 638)

AUTHORS Mahatras, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T., Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and Hood, L.

TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

COMMENT Contact: Mahatras GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RPCI-11. For BAC library availability, please contact Pieter de Jong (pieter@jlong.med.buffalo.edu). Clones may be purchased from BACPAC Resources (<http://bacpac.med.buffalo.edu/ordering/bac.htm>) or from Research Genetics (<http://www.htsc.washington.edu>)

Plates: 1112 row: D column: 19

Seq primer: T7

Class: BAC ends

High quality sequence stop: 638.

Location/Qualifiers

1..638

/organism="Homo sapiens"

/mol\_type="Genomic DNA"

/db\_xref="taxon:9606"

/clone="Plate=1112 Col=19 Row=D"

/sex="male"

/clone\_1lb="RPCI-11 Human Male BAC library"

/note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"

ORIGIN

Query Match 1.4%; Score 44; DB 9; Length 638;

Best Local Similarity 100.0%; Pred. No. 1.9e-09;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TGCCAGTGCAGCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122

DB 353 TGCCAGTGCAGCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 310

RESULT 171

LOCUS AV733586 641 bp mRNA linear EST 17-OCT-2000

DEFINITION AV733586 cda Homo sapiens cDNA clone cdaACG10 5', mRNA sequence.

ACCESSION AV733586

VERSION AV733586.1 GI:10851131

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 641)

AUTHORS Yang, Y., Song, H., Peng, Y., Gu, Y., Gao, G., Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Ou, J., Gao, X., Cheng, Z., Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z., and Han, Z.

Human sapiens cDNA clones

Unpublished (2000)

CONTACT: Zeguang Han

Chinese National Human Genome Center at Shanghai

351 Guo Shouling Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai 201203, P. R. China

Tel: 86-21-50801919 (ex. 45)

Fax: 86-21-50801922

Email: hanzg@chgc.sh.cn

This clone is available at CHGC in Shanghai.

Location/Qualifiers

```
source
1..641
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="cdACG10"
/cisue_type="pneocromocytoma"
/dev_stage="Adult"
/lab_host="BM25.8"
/clone_lib="cda"
/note="vector: pTriplEx2; Site_1: sflrA; Site_2: sflrB"

ORIGIN
Query Match      1.4%; Score 44; DB 1; Length 641;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3073 AGATTGTGCGCACTGCATCCAGCCTGGGCAACAGACAGACTC 3116
        |||||||
Db      406 AGATTGTGCGCACTGCATCCAGCCTGGGCAACAGACAGACTC 449

RESULT 172
AG025905      649 bp      DNA      linear      GSS 16-FEB-2005
LOCUS      AG025905/C
DEFINITION      Homo sapiens genomic DNA, 21q region, clone: B335N5 A043 (Fw),
GENOMIC SURVEY SEQUENCE.
ACCESSION      AG025905
VERSION      AG025905.1 GI:6594369
KEYWORDS      GSS.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE      1
AUTHORS      Hattori,M., Fujiyama,A., Ishii,K., Toyoda,A., Taylor,T.,
              Park,H.-S., Yada,T., Watanabe,H. and Sakaki,Y.
TITLE      Homo sapiens genomic DNA, chromosome 21q
JOURNAL      Published Only in Database (1999)
AUTHORS      2 (bases 1 to 649)
              Hattori,M., Fujiyama,A., Ishii,K., Toyoda,A., Taylor,T.,
              Park,H.-S., Yada,T., Watanabe,H. and Sakaki,Y.
TITLE      Direct Submission
JOURNAL      Submitted (17-DEC-1999) Masahira Hattori, RIKEN Genomic Sciences
              Center, RIKEN Yokohama Institute, Yokohama Research Promotion
              Division, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa,
              230-0045, Japan (E-mail:hattori@gs.c.riken.jp, Tel:81-45-503-9111,
              Fax:81-45-503-9113)
FEATURES
source
1..649
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="21"
/map="21q"
/clone="B335N5 A043 (Fw)"

ORIGIN
Query Match      1.4%; Score 44; DB 10; Length 649;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2889 GAGGCGAGTGATCAGCTGAGGCGAGAGTGTGAGACGAGCTG 2932
        |||||||
Db      389 GAGGCGAGTGATCAGCTGAGGCGAGAGTGTGAGACGAGCTG 346

RESULT 173
AG080818      650 bp      DNA      linear      GSS 03-NOV-2001
LOCUS      AG080818
DEFINITION      Pan troglodytes DNA, clone: PTB-077B13.R, genomic survey sequence.
ACCESSION      AG080818
VERSION      AG080818.1 GI:16632620
```

```
KEYWORDS      GSS.
SOURCE      Pan troglodytes (chimpanzee)
ORGANISM      Pan troglodytes
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Pan.
REFERENCE      1
AUTHORS      Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
              Torokl,Y., Watanabe,H. and Sakaki,Y.
TITLE      BAC end sequences of Library PTB
JOURNAL      Unpublished
AUTHORS      2 (bases 1 to 650)
              Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
              Torokl,Y., Watanabe,H. and Sakaki,Y.
TITLE      Direct Submission
JOURNAL      Submitted (02-AUG-2001) Aseo Fujiyama, The Institute of Physical
              and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Japan
              1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
              (E-mail:chimbee@gs.c.riken.go.jp, URL:http://hgp-gsc.riken.go.jp/,
              Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT      Clones are derived from the chimpanzee BAC library PTB. This BAC end
              was generated during the Rad process and may have higher chance of
              clone tracking errors.
              PRIMERS
              Sequencing: M13Rev
LIBRARY      Vector : pKS145
              R.Site 1 : SacI
              R.Site 2 : SacI.
              Location/Qualifiers
FEATURES
source
1..650
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"
/clone="PTB-077B13.R"
/sex="male"
/cell_type="lymphoblast"
/clone_lib="PTB Chimpanzee Male BAC Library"

ORIGIN
Query Match      1.4%; Score 44; DB 10; Length 650;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3073 AGATTGTGCGCACTGCATCCAGCCTGGGCAACAGACAGACTC 3116
        |||||||
Db      221 AGATTGTGCGCACTGCATCCAGCCTGGGCAACAGACAGACTC 264

RESULT 174
AV722027      652 bp      mRNA      linear      EST 16-OCT-2000
LOCUS      AV722027
DEFINITION      HTB Homo sapiens cDNA clone HTBBZD11.5', mRNA sequence.
ACCESSION      AV722027
VERSION      AV722027.1 GI:10824105
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Hominiidae; Homo.
REFERENCE      1 (bases 1 to 652)
AUTHORS      Gu,Y., Peng,Y., Song,H., Huang,Q., Yang,Y., Gao,G., Xiao,H., Xu,X.,
              Li,N., Qian,B., Liu,F., Qu,J., Gao,X., Cheng,Z., Xu,Z., Zeng,L.,
              Xu,S., Gu,W., Tu,Y., Jia,J., Fu,G., Ren,S., Zhong,M., Lu,G., Hu,R.,
              Chen,J., Chen,Z. and Han,Z.
TITLE      Homo sapiens cDNA HTB clones
JOURNAL      Unpublished (2000)
COMMENT      Contact: Zeguang Han
              Chinese National Human Genome Center at Shanghai
              351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
              201203, P. R. China
              Tel: 86-21-50801919(ex.45)
```

FEATURES  
source  
Fax: 86-21-50801922  
Email: hanzg@chgc.sh.cn  
This clone is available at CHGC in Shanghai.  
Location/Qualifiers

1.652  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="HTB2D11"  
/issue\_type="Hypothalamus"  
/dev\_stage="Adult"  
/lab\_host="SOLR"  
/clone\_1lb="HTB"  
/note="Vector: pBluescript sk(-); Site\_1: EcoRI; Site\_2:  
XhoI"

## ORIGIN

Query Match 1.4%; Score 44; DB 1; Length 652;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TGCCACTGCATCCAGCTCGGCAACAGCAAGACTCTGTCTC 3122  
Db 25 TGCCACTGCATCCAGCTCGGCAACAGCAAGACTCTGTCTC 68

RESULT 175  
AG106557/c 654 bp DNA linear GSS 03-NOV-2001  
LOCUS Pan troglodytes DNA, clone: PTB-111B19.F, genomic survey sequence.  
ACCESSION AG106557  
VERSION AG106557.1 GI:16727075  
KEYWORDS GSS.  
SOURCE Pan troglodytes (chimpanzee)  
ORGANISM Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Pan.

REFERENCE 1  
AUTHORS Fujiyama, A., Hatori, M., Toyoda, A., Taylor, T.D., Yada, T.,  
Toto, Y., Watanabe, H. and Sakaki, Y.  
TITLE BAC end sequences of library PTB  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 654)  
AUTHORS Fujiyama, A., Hatori, M., Toyoda, A., Taylor, T.D., Yada, T.,  
Toto, Y., Watanabe, H. and Sakaki, Y.  
TITLE Direct Submission  
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);  
1-7-22 Shuhiro-chou, Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail: chihiro@gsc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/  
Tel: 81-45-503-9111, Fax: 81-45-503-9170)

COMMENT Clones are derived from the chimpanzee BAC library PTB this BAC end  
was generated during the R&D process and may have higher chance of  
clone tracking errors.  
PRIMERS  
Sequencing: -21M13

LIBRARY Vector : pKS145  
R.Site 1 : SacI  
R.Site 2 : SacI.  
Location/Qualifiers  
1.654  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone="PTB-111B19.F"  
/sex="male"  
/cell\_type="lymphoblast"  
/clone\_1lb="PTB Chimpanzee Male BAC library"

## FEATURES

source

## ORIGIN

Query Match 1.4%; Score 44; DB 10; Length 654;

Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGCGCATCTCCAGCTCGGCAACAGCAAGACTCTGTCTC 3116  
Db 370 AGATTGCGCATCTCCAGCTCGGCAACAGCAAGACTCTGTCTC 327

## RESULT 176

AW959999

LOCUS AW959999 655 bp mRNA linear EST 01-JUN-2000  
DEFINITION EST372070 MAGF resequences, MAGF Homo sapiens cDNA, mRNA sequence.  
ACCESSION AW959999  
VERSION AW959999.1 GI:8149683  
KEYWORDS EST.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.

REFERENCE 1 (bases 1 to 655)  
AUTHORS Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C.,  
Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeaman, T.J. and  
Quackenbush, J.  
TITLE Assessment of gene expression patterns in a model of colon tumor  
metastasis using a 19,200 element cDNA microarray  
JOURNAL Unpublished (2000)

CONTACT: John Quackenbush  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 3528  
Fax: 301 838 0208  
Email: john@eigr.org  
Plate: 140  
Seq primer: Reverse.

FEATURES  
source  
1.655  
Location/Qualifiers

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_1lb="MAGF resequences, MAGF"  
/note="Vector: pBluescriptSM"

## ORIGIN

Query Match 1.4%; Score 44; DB 1; Length 655;  
Best Local Similarity 100.0%; Pred. No. 1.9e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TGCCACTGCATCCAGCTCGGCAACAGCAAGACTCTGTCTC 3122  
Db 187 TGCCACTGCATCCAGCTCGGCAACAGCAAGACTCTGTCTC 230

## RESULT 177

AO740363

LOCUS AO740363 694 bp DNA linear GSS 16-JUL-1999  
DEFINITION HS\_5501\_A1\_G11 SR6 RPCI-11 Human Male BAC library Homo sapiens  
genomic clone Plate=1077 Col=21 Row=W, genomic survey sequence.  
ACCESSION AO740363  
VERSION AO740363.1 GI:5517885  
KEYWORDS GSS.

SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.

REFERENCE 1 (bases 1 to 694)  
AUTHORS Mahairas, G.G., Wallace, J.C., Smith, K., Swartzell, S., Holzman, T.,  
Keller, A., Shaker, R., Furlong, J., Young, J., Zhao, S., Adams, M.D. and  
Hood, L.  
TITLE Sequence-tagged connectors: A sequence approach to mapping and  
scanning the human genome  
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

Query Match 1.4%; Score 44; DB 10; Length 654;

PUBMED  
10449764  
Contact: Mahairas GC, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Queen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones are derived from the human BAC library RPC1-11. For BAC  
library availability, please contact Pletcher de Jong  
(pletched@u.washington.edu). Clones may be purchased from  
BACpac Resources (<http://bacpac.med.buffalo.edu/ordering/bac.htm>)  
or from Research Genetics ([info@resgen.com](http://info@resgen.com)). BAC end Web Server:  
<http://www.htsc.washington.edu>  
Place: 1077 Row: M Column: 21  
Seq primer: SP6  
Class: BAC ends  
High quality sequence stop: 694.  
Location/Qualifiers  
1..694  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="Place=1077 Col=21 Row=M"  
/sex="male"  
/clone\_lib="RPC1-11 Human Male BAC Library"  
/note="Vector: pBAC3.6; Site\_1: EcoRI; Site\_2: EcoRI;  
Male blood DNA was isolated from one randomly chosen donor  
and partially digested with a combination of EcoRI and  
EcoRI Methylase. Size selected DNA was cloned into the  
pBAC3.6 vector at EcoRI sites"

ORIGIN  
Query Match 1.4%; Score 44; DB 9; Length 694;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TGGCAGTGCATCCAGCTGGGCAACAGCAAGACTGTCTC 3122  
|||||  
156 TGGCAGTGCATCCAGCTGGGCAACAGCAAGACTGTCTC 199  
|||||

RESULT 178  
LOCUS BU616112/c 696 bp mRNA linear EST 23-SEP-2002  
DEFINITION UI-H-DFO-bex-m-18-0-UI.s1 NCI CGAP DFO Homo sapiens cDNA clone  
BU616112  
ACCESSION BU616112  
VERSION BU616112.1 GI:23282320  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 696)  
NCI-CCG <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: cgapbs-r@mail.nih.gov  
Tissue Procurement: Dr. Jose Mercuende  
cDNA library preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained  
from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
The following repetitive elements were found in this cDNA  
sequence: 11-299, >ALU (matched complement) 647-696, >ALU  
Seq primer: M13 FORWARD  
POLYA=yes.  
Location/Qualifiers

FEATURES  
source

source  
1..696  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-H-DFO-bex-m-18-0-UI"  
/issue\_type="Subchondral Bone"  
/dev\_stage="Adult"  
/lab\_host="DHI08 (Life Technologies)"  
/clone\_lib="NCI CGAP DFO"  
/note="Organ: Bone; Vector: pRTT3-Pac (Pharmacia) with a  
modified polylinker; Site 1: EcoRI; Site 2: Not I;  
NCI CGAP DFO is a cDNA library containing the following  
tissue(s): Subchondral Bone. The library was constructed  
according to Bonaldo, Lennon and Soares, Genome Research,  
6:791-806, 1996. First strand cDNA synthesis was primed  
with an oligo-dT primer containing a Not I site. Double  
stranded cDNA was ligated to an EcoRI adaptor, digested  
with Not I, and cloned directionally into pRTT3-Pac  
vector. The oligonucleotide used to prime the synthesis of  
first-strand cDNA contains a library tag sequence that is  
located between the Not I site and the (dT)18 tail. The  
sequence tag for this library is GTTAAGCGTC.  
TAG TISSUE=Subchondral bone  
TAG LIB=UI-H-DFO  
TAG\_SEQ=GTTAAGCGTC"

ORIGIN  
Query Match 1.4%; Score 44; DB 5; Length 696;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGGCAGTGCATCCAGCTGGGCAACAGCAAGACTC 3116  
|||||  
Db 68 AGATTGGCAGTGCATCCAGCTGGGCAACAGCAAGACTC 25  
|||||

RESULT 179  
LOCUS AQ262202 706 bp DNA linear GSS 24-OCT-1998  
DEFINITION CITR1-B1-2509J18.TF CITR1-B1 Homo sapiens genomic clone 2509J18,  
genomic survey sequence.  
ACCESSION AQ262202  
VERSION AQ262202.1 GI:3788682  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 706)  
Adams,M.D., Rounsley,S.D., Zhao,S., Bass,S., Linher,K., Golden,K.,  
Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H., Simon,M. and  
Venter,J.C.  
Use of a random human BAC End Sequence Database for Sequence-Ready  
Map Building  
Unpublished (1998)  
Other\_GSSes: CITR1-B1-2509J18.TR  
Contact: Mark Adams  
Department of Eukaryotic Genomics  
The Institute for Genomic Research  
9712 Medical Center Dr., Rockville, MD 20850, USA  
Tel: 301 838 0200  
Fax: 301 838 0208  
Email: mdamad@tigr.org  
Clones are available from Research Genetics ([info@resgen.com](http://info@resgen.com)). BAC  
end search page:  
[http://www.tigr.org/cdb/hungen/bac\\_end\\_search/bac\\_end\\_search.html](http://www.tigr.org/cdb/hungen/bac_end_search/bac_end_search.html).  
Seq primer: M13-21  
Class: BAC ends.  
Location/Qualifiers  
1..706  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"

FEATURES  
source



/db\_xref="taxon:9606"  
/clone="2509J18"  
/sex="male"  
/cell\_type="sperm"  
/clone\_lib="CITBI-81"  
/note="Vector: pBelBAC11; Site\_1: EcoRI; Site\_2: EcoRI;  
Caltech Human BAC Library D"

## ORIGIN

Query Match 1.4%; Score 44; DB 9; Length 706;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3072 AGATTGTGCGCACTGCATCCAGCTCGGGCAACAGAGCAAGACTC 3116  
|||||  
DB 538 AGATTGTGCGCACTGCATCCAGCTCGGGCAACAGAGCAAGACTC 581

RESULT 180  
BX504396 715 bp mRNA linear EST 04-SEP-2003  
LOCUS DKFZP686O22133.r1.686 (synonym: h1ccc3) Homo sapiens cDNA clone  
DEFINITION DKFZP686O22133 5', mRNA sequence.  
ACCESSION BX504396  
VERSION BX504396.1 GI:32031347  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
REFERENCE 1 (bases 1 to 715)  
AUTHORS Poustka, A., Alber, R., Moosmayer, P., Schnupp, I., Wellenreuther, R.,  
Wewes, H. W., Well, B., Amlid, C., Obanger, A., Fobo, G., Han, M. and  
Wiemann, S.  
TITLE EST (Poustka, A., Alber, R., Moosmayer, P., Schnupp, I.,  
Wellenreuther, R., et al.)  
JOURNAL Unpublished (2003)  
COMMENT Contact: MTPS

MFPS  
Ingelstaedter Landstr. 1, D-85764 Neuherberg, Germany  
This is the 5' sequence of the clone insert  
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer  
Research Center (DKFZ), Email: s.wiemann@dkfz-heidelberg.de;  
Heidelberg/Germany) within the cDNA sequencing consortium of the  
German Genome Project.  
No sl sequence available.  
This clone (DKFZP686O22133) is available at the RZPD in Berlin.  
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059  
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.  
Location/Qualifiers

## FEATURES

source

1..715  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="DKFZP686O22133"  
/dev\_stage="adult"  
/lab\_host="DH10B"  
/clone\_lib="686 (synonym: h1ccc3)"  
/note="Vector: pTriplEx2; site\_1: SfiI; site\_2: SfiI;  
cDNA-collection"

## ORIGIN

Query Match 1.4%; Score 44; DB 5; Length 715;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3072 AGATTGTGCGCACTGCATCCAGCTCGGGCAACAGAGCAAGACT 3115  
|||||  
DB 255 AGATTGTGCGCACTGCATCCAGCTCGGGCAACAGAGCAAGACT 298

RESULT 181  
BX098422/c 716 bp mRNA linear EST 04-FEB-2003  
LOCUS BX098422 Soares fetal liver spleen INFLS Homo sapiens cDNA clone  
DEFINITION IMAGP98J0883 ; IMAGE:109927, mRNA sequence.  
ACCESSION BX098422  
VERSION BX098422.1 GI:27829399  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homidae; Homo.  
REFERENCE 1 (bases 1 to 716)  
AUTHORS Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M.,  
Radolf, U., Schneider, D. and Korn, B.  
TITLE Human Unigeneset - RZPD3  
JOURNAL Unpublished (2003)  
COMMENT Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany  
RZPD; IMAGP98J0883.  
RZPDLIB; I.M.A.G.-E. cDNA Clone Collection;  
Human Unigeneset - RZPD3 (RZPDLIB No. 972)  
http://www.rzpd.de/cloneCards/cgi-bin/showLib.pl.cgi?response?libNo=972 Contact: Ina Rolfs  
RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
Heubnerweg 6, D-14059 Berlin, Germany  
Tel: +49 30 32639 101  
Fax: +49 30 32639 111  
www.rzpd.de

FEATURES  
source  
This clone is available royalty-free from RZPD;  
contact RZPD (clone@rzpd.de) for further information. Seq primer:  
M13r, Primer sequence: TTTCACAGGAAACAGCTATGAC.  
Location/Qualifiers

1..716  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGP98J0883 ; IMAGE:109927"  
/sex="male"  
/dev\_stage="20 week-post conception fetus"  
/lab\_host="DH10B (ampicillin resistant)"  
/clone\_lib="Soares fetal liver spleen INFLS"  
/note="Organ: Liver and Spleen; Vector: pTriplEx2 (Pharmacia)  
with a modified polylinker; Site\_1: Pac I; Site\_2: Eco RI;  
1st strand cDNA was primed with a Pac I - oligo(dT) primer  
[5' AACTGGAAGATTATTAAGATCTTTTCTTTTCTTTT 3']  
double-stranded cDNA was ligated to Eco RI adaptors  
(Pharmacia), digested with Pac I and cloned into the Pac I  
and Eco RI sites of the modified pTriplEx2 vector. Library  
went through one round of normalization. Library  
constructed by Bento Soares and M. Fatima Bernaldo."

## ORIGIN

Query Match 1.4%; Score 44; DB 5; Length 716;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TCCGCTGCACTCCAGCTCGGGCAACAGAGCAAGCTGTCTC 3122  
|||||  
DB 140 TCCGCTGCACTCCAGCTCGGGCAACAGAGCAAGCTGTCTC 97

RESULT 182  
AG175100/c 745 bp DNA linear GSS 09-JAN-2002  
LOCUS AG175100  
DEFINITION Pan troglodytes DNA, clone: RP43-045024.TJ, genomic survey  
sequence.  
ACCESSION AG175100  
VERSION AG175100.1 GI:16704780  
KEYWORDS GSS.  
SOURCE Pan troglodytes (chimpanzee)

ORGANISM Pan troglodytes  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Pan.

REFERENCE 1  
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,  
Toto, Y., Matanabe, H. and Sakaki, Y.  
TITLE BAC end sequences of library RPCI-43  
JOURNAL Unpublished  
AUTHORS 2 (bases 1 to 745)  
TITLE Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T. D., Yada, T.,  
Toto, Y., Matanabe, H. and Sakaki, Y.  
JOURNAL Direct Submission  
SUBMITTED (02-AUG-2001) Asao Fujiyama, The Institute of Physical  
and Chemical Research (RIKEN), Genomic Sciences Center (GSC), Japan  
1-7-22, Suenitro-chou, Tsunumi-Ku, Yokohama, Kanagawa 230-0045, Japan  
(E-mail:chimbases@sc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,  
Tel:81-45-503-9111, Fax:81-45-503-9170)  
Clones are derived from the chimpanzee BAC library RPCI-43 This BAC  
end was generated during the R&D process and may have higher chance  
of clone tracking errors.  
PRIMERS

COMMENT Sequencing: TJ  
LIBRARY  
Vector : pBAC3.6  
R.Site 1 : EcoRI  
R.Site 2 : EcoRI.  
Location/Qualifiers  
1..745  
/organism="Pan troglodytes"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9598"  
/clone="RP43-045024.TJ"  
/sex="male"  
/cell\_type="lymphocytes"  
/clone\_1lb="RPCI-43 Chimpanzee Male BAC library"

ORIGIN

Query Match 1.4%; Score 44; DB 10; Length 745;  
Best Local Similarity 100.0%; Pred.No. 1.8e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 GTGGCACTGCATCTCCAGCTGGGCAACAGAGAAGACTGTGCT 3121  
|||||  
Db 478 GTGGCACTGCATCTCCAGCTGGGCAACAGAGAAGACTGTGCT 435

RESULT 183  
CB308313 781 bp mRNA linear EST 04-MAR-2003  
LOCUS CB308313  
DEFINITION AGENCOURT 11818308 NICHD\_Rh\_Ov1 Macaca mulatta cDNA clone  
IMAGE:6682307 5', mRNA sequence.  
ACCESSION CB308313  
VERSION CB308313.1 GI:28831023  
KEYWORDS EST.  
SOURCE Macaca mulatta (rhesus monkey)  
ORGANISM Macaca mulatta  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Cercopitheciidae; Cercopitheciinae; Macaca.  
1 (bases 1 to 781)  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
Tissue Procurement: Dr. Blot Spindel  
cDNA Library Preparation: Clontech  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LIML at:

COMMENT JOURNAL  
AUTHORS  
TITLE  
JOURNAL  
UNPUBLISHED (1997)  
CONTACT: Robert Strausberg, Ph.D.  
EMAIL: [cgabbs-r@mail.nih.gov](mailto:cgabbs-r@mail.nih.gov)  
TISSUE PROCUREMENT: Dr. Blot Spindel  
CDNA LIBRARY PREPARATION: Clontech  
CDNA LIBRARY ARRAYED BY: THE I.M.A.G.E. CONSORTIUM (LIML)  
DNA SEQUENCING BY: AGENCOURT BIOSCIENCE CORPORATION  
CLONE DISTRIBUTION: NCI-CGAP CLONE DISTRIBUTION INFORMATION CAN BE  
FOUND THROUGH THE I.M.A.G.E. CONSORTIUM/LIML AT:

<http://image.llnl.gov>  
Plate: LNCM3128 row: 1 column: 18  
High quality sequence spot: 544.  
Location/Qualifiers  
1..781  
/organism="Macaca mulatta"  
/mol\_type="mRNA"  
/db\_xref="taxon:9544"  
/clone="IMAGE:6682307"  
/rissue\_type="Ovary"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_1lb="NICHD\_Rh\_Ov1"  
/note="Organ: ovary; Vector: pDNR-LIB; Site 1: Sfi I;  
Site 2: Sfi I; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.0-4.0 kb. Tissue pooled from  
pre-puberatal, post pubertal an menopausal monkeys.  
Constructed by Clontech. Note: this is a NICHD library."

ORIGIN

Query Match 1.4%; Score 44; DB 6; Length 781;  
Best Local Similarity 100.0%; Pred.No. 1.8e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3079 TGCCACTGCATCTCCAGCTGGGCAACAGAGAAGACTGTGCTC 3122  
|||||  
Db 205 TGCCACTGCATCTCCAGCTGGGCAACAGAGAAGACTGTGCTC 248

RESULT 184  
B2610701 803 bp DNA linear GSS 08-JUN-2003  
LOCUS B2610701/C  
DEFINITION WHAMJ70TR Human MCF7 breast cancer cell line library (MCF7\_1) Homo  
sapiens genomic clone MCF7\_1-3K20, genomic survey sequence.  
ACCESSION B2610701  
VERSION B2610701.1 GI:31519262  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Homnidae; Homo.  
1 (bases 1 to 803)  
Volik, S., Zhao, S., Chin, K., Brebner, J.H., Herndon, D.R., Tao, Q.,  
Kowbel, D., Huang, G., Lapuk, A., Kuo, W.-L., Magrane, G., de Jong, P.,  
Gray, J.W. and Collins, C.  
End-sequence profiling: Sequence-based analysis of aberrant genomes  
Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7656-7701 (2003)  
12788976  
Contact: Volik SV  
Colin Collins' lab  
UCSF Comprehensive Cancer Center  
UCSF Box 0808, San Francisco, CA 94143-0808, USA  
Tel: 415 502 7066  
Fax: 415 502 5665  
Email: [svolik@cc.ucsf.edu](mailto:svolik@cc.ucsf.edu)  
This clone is available from Amplicon Express  
<http://www.genomex.com>  
Class: BAC ends.

FEATURES  
source  
1..803  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="MCF7\_1-3K20"  
/sex="female"  
/clone\_1lb="Human MCF7 breast cancer cell line library  
(MCF7\_1)"  
/note="Vector: pECBAC1; Site 1: HindIII; This library was  
constructed from MCF7 breast cancer cell line by Amplicon  
Express (<http://www.genomex.com>) using their standard  
procedure."

ORIGIN

Query Match 1.4%; Score 44; DB 9; Length 803;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACTGCATCTCCAGCTTGCGCAACAGCAAGACTC 3116  
DB 221 AGATTGTGCACTGCATCTCCAGCTTGCGCAACAGCAAGACTC 178

## RESULT 185

CA442440 828 bp mRNA linear EST 08-NOV-2002  
LOCUS UI-H-DPO-ave-m-01-0-UI-61 NCI CGAP F81 Homo sapiens cDNA clone  
DEFINITION UI-H-DPO-ave-m-01-0-UI 3', mRNA sequence.  
ACCESSION CA442440.1 GI:24806860  
VERSION CA442440.1 GI:24806860  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

## REFERENCE

1 (bases 1 to 828)  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)

## JOURNAL

CONTACT: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Dr. Mary Hendrix  
CDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
Clone Distribution: Clone distribution information can be obtained  
from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
The following repetitive elements were found in this cDNA  
sequence: 163-455, >ALU 416-466, >ALU  
Seq primer: M13 FORWARD  
POLYA=yes.

## COMMENT

## FEATURES

source Location/Qualifiers  
1..828

/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="UI-H-DPO-ave-m-01-0-UI"  
/tissue\_type="Fibroblasts"  
/lab\_host="DH10B (Life Technologies)"  
/clone\_1ib="NCI CGAP F81"  
/note="Vector: pTR73-Pac (Pharmacia) with a modified  
polylinker; Site 1: EcoR I; Site 2: Not I; NCI CGAP\_F81 is  
a cDNA library containing the following tissue(s):  
Fibroblasts Cell line HT-1080 (ATCC number CCL-121). The  
library was constructed according to Bonaldi, Lennon and  
Soares, Genome Research, 6:791-806, 1996. First strand  
cDNA synthesis was primed with an oligo-dT primer  
containing a Not I site. Double stranded cDNA was ligated  
to an EcoR I adaptor, digested with Not I, and cloned  
directionally into pTR73-Pac vector. The oligonucleotide  
used to prime the synthesis of first-strand cDNA contains  
a library tag sequence that is located between the Not I  
site and the (dT)18 tail. The sequence tag for this  
library is GTCTACGAG.  
TAG\_L1B=UI-H-DPO  
TAG\_SEQ=GTCTACGAG"

## ORIGIN

Query Match 1.4%; Score 44; DB 6; Length 828;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACTGCATCTCCAGCTTGCGCAACAGCAAGACTC 3116  
|||||

DB 398 AGATTGTGCACTGCATCTCCAGCTTGCGCAACAGCAAGACTC 441

## RESULT 186

CB307968 830 bp mRNA linear EST 04-MAR-2003  
LOCUS AGENCOURT 11771960 NICHHD\_Rh\_OV1 Macaca mulatta cDNA clone  
DEFINITION IMAGE:6882266 5', mRNA sequence.  
ACCESSION CB307968.1 GI:28830678  
VERSION CB307968.1 GI:28830678  
KEYWORDS EST.  
SOURCE Macaca mulatta (rhesus monkey)  
ORGANISM Macaca mulatta

## REFERENCE

1 (bases 1 to 830)  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Cercopithecoidea; Cercopithecoidea; Macaca.  
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
Tumor Gene Index  
Unpublished (1997)

## JOURNAL

CONTACT: Robert Strausberg, Ph.D.  
Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
Tissue Procurement: Dr. Eliot Spindel  
CDNA Library Preparation: CLOUTCH  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone Distribution: NCI-CGAP clone distribution information can be  
found through the I.M.A.G.E. Consortium/LLNL at:  
<http://image.llnl.gov>  
Plate: L10C3128 row: k column: 01  
High quality sequence stop: 499.

## COMMENT

## FEATURES

source Location/Qualifiers  
1..830  
/organism="Macaca mulatta"  
/mol\_type="mRNA"  
/db\_xref="taxon:9544"  
/clone="IMAGE:6882266"  
/tissue\_type="Ovary"  
/lab\_host="DH10B (phage-resistant)"  
/clone\_1ib="NICHHD\_Rh\_OV1"  
/note="Organ: ovary; Vector: pDNR-LIB; Site 1: Sfi I;  
Site 2: Sfi I; Cloned unidirectionally. Primer: Oligo dT.  
Average insert size 1.0-4.0 kb. Tissue pooled from  
pre-pubertal, post pubertal and menopausal monkeys.  
Constructed by Clontech. Note: this is a NICHHD library."

## ORIGIN

Query Match 1.4%; Score 44; DB 6; Length 830;  
Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACTGCATCTCCAGCTTGCGCAACAGCAAGACTC 3116  
DB 365 AGATTGTGCACTGCATCTCCAGCTTGCGCAACAGCAAGACTC 322

## RESULT 187

A0781442 830 bp DNA linear GSS 02-AUG-1999  
LOCUS A0781442/c HS 3117 A2 G10 TTC CIT Approved Human Genomic Sperm Library D Homo  
DEFINITION sapiens genomic clone Plate=3117 Col=20 Row=M, genomic survey  
sequence.  
ACCESSION A0781442  
VERSION A0781442.1 GI:5684402  
KEYWORDS GSS.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

## REFERENCE

1 (bases 1 to 830)  
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;  
Hominidae; Homo.  
Mahatras,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,  
REFERENCES

```
Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
Hood,L.
Sequence-tagged connectors: A sequence approach to mapping and
scanning the human genome
Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
10449764
Contact: Mahairas GG, Wallace JC, Hood L
High Throughput Sequencing Center
University of Washington
401 Queen Anne Avenue North, Seattle, WA 98109, USA
Tel: (206) 616-3618
Fax: (206) 616-3887
Email: jwallace@u.washington.edu
Clones may be purchased from Research Genetics (info@resgen.com).
BAC end Web Server: http://www.htec.washington.edu
Plate: 317 Row: M Column: 20
Seq primer: T7
Class: BAC ends
High quality sequence stop: 830.

FEATURES
    source
        1..830
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /clone="MCF7_1-22L24"
        /sex="female"
        /clone_1lib="Human MCF7 breast cancer cell line library
        (MCF7_1)"
        /note="Vector: pECBAC1, Site 1: HindIII; This library was
        constructed from MCF7 breast cancer cell line by Amplicon
        Express (http://www.genomex.com) using their standard
        procedure."

ORIGIN
Query Match          1.4%; Score 44; DB 9; Length 834;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3079  TGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGCTCTGTCTC 3122
DB      204  TGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGCTCTGTCTC 161

RESULT 189
LOCUS      A0781550      847 bp      DNA      linear      GSS 02-AUG-1999
DEFINITION HS 3176 A2 A08 T7C CIT Approved Human Genomic Sperm Library D Homo
            sapiens genomic clone Plate=3176 Col=16 Row=A, genomic survey
            sequence.
ACCESSION  A0781550
VERSION     A0781550.1 GI:5684510
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens (human)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo.
            1 (bases 1 to 847)
            Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holman,T.,
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
            Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
            Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
            10449764
            Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
            Clones may be purchased from Research Genetics (info@resgen.com).
            BAC end Web Server: http://www.htec.washington.edu
            Plate: 3176 Row: A Column: 16
            Seq primer: T7
            Class: BAC ends
            High quality sequence stop: 847.

FEATURES
    source
        1..847
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /clone="Plate=3176 Col=16 Row=A"
        /sex="male"
        /clone_1lib="CIT Approved Human Genomic Sperm Library D"
        /note="Organ: sperm; Vector: pBelBAC11; BAC Clones in
        B-Col1 DH10B"

ORIGIN
Query Match          1.4%; Score 44; DB 9; Length 847;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

2895  GGTGATGACCTGAGCGCAGAGTTCCAGACCAAGCTGCGCAAC 2938
DB      2895  GGTGATGACCTGAGCGCAGAGTTCCAGACCAAGCTGCGCAAC 2938

RESULT 188
LOCUS      BZ612098      834 bp      DNA      linear      GSS 08-JUN-2003
DEFINITION WHABP72TR Human MCF7 breast cancer cell line library (MCF7_1) Homo
            sapiens genomic clone MCF7_1-22L24, genomic survey sequence.
ACCESSION  BZ612098
VERSION     BZ612098.1 GI:31520659
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens (human)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo.
            1 (bases 1 to 834)
            Volk,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,
            Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
            Gray,J.W. and Collins,C.
            Bnd-sequence profiling: Sequence-based analysis of aberrant genomes
            Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
            12788976
            Contact: Volk SV
            Colln Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7066
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            This clone is available from Amplicon Express
            http://www.genomex.com
            Class: BAC ends
            Location/Qualifiers
                1..834
                /organism="Homo sapiens"

ORIGIN
Query Match          1.4%; Score 44; DB 9; Length 830;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3072  AAGATTGTGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGCT 3115
DB      454  AAGATTGTGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGCT 411

RESULT 188
LOCUS      BZ612098      834 bp      DNA      linear      GSS 08-JUN-2003
DEFINITION WHABP72TR Human MCF7 breast cancer cell line library (MCF7_1) Homo
            sapiens genomic clone MCF7_1-22L24, genomic survey sequence.
ACCESSION  BZ612098
VERSION     BZ612098.1 GI:31520659
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens (human)
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homiidae; Homo.
            1 (bases 1 to 834)
            Volk,S., Zhao,S., Chin,K., Brebner,J.H., Herndon,D.R., Tao,Q.,
            Kowbel,D., Huang,G., Lapuk,A., Kuo,W.-L., Magrane,G., de Jong,P.,
            Gray,J.W. and Collins,C.
            Bnd-sequence profiling: Sequence-based analysis of aberrant genomes
            Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
            12788976
            Contact: Volk SV
            Colln Collins' lab
            UCSF Comprehensive Cancer Center
            UCSF Box 0808, San Francisco, CA 94143-0808, USA
            Tel: 415 502 7066
            Fax: 415 502 5665
            Email: svolik@cc.ucsf.edu
            This clone is available from Amplicon Express
            http://www.genomex.com
            Class: BAC ends
            Location/Qualifiers
                1..834
                /organism="Homo sapiens"

ORIGIN
Query Match          1.4%; Score 44; DB 9; Length 830;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3072  AAGATTGTGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGCT 3115
DB      454  AAGATTGTGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGCT 411
```

Db 322 GGTGATCATCCTGAGCCGAGGATTGAGACCAAGCTGGCCAC 279

## RESULT 190

BF573285

867 bp mRNA linear EST 12-DEC-2000

LOCUS 602079678F2\_NTH\_MGC\_62 Homo sapiens cDNA clone IMAGE:4254115 5', mRNA sequence.

## ACCESSION

BF573285

BF573285

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

## FEATURES

source

High quality sequence stop: 290.

Location/Qualifiers

1..867  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:4254115"  
/cfeature\_type="melanotic melanoma, high MDR"  
/lab\_host="DH10B (T1 phage-resistant)"  
/clone\_1ib="NTH MGC\_62"  
/note="Organ: skin; Vector: pDR-LIB (Clontech); site\_1: SfiI (ggcgccgcgcgcgc); site\_2: SfiI (ggcgccgcgcgc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-ATCTAGAGCGCGCGCCGACATG-3' (30) BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.75 kb (range 0.9-4.0 kb). 15/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

## ORIGIN

Query Match 1.4%; Score 44; DB 2; Length 867;

Best Local Similarity 100.0%; Pred. No. 1.8e-09;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3073 AGATTGTCACCTGACCTCCAGCCTGGGCAACAGAGCAAGCTC 3116

228 AGATTGTCACCTGACCTCCAGCCTGGGCAACAGAGCAAGCTC 271

Db 228 AGATTGTCACCTGACCTCCAGCCTGGGCAACAGAGCAAGCTC 271

## RESULT 191

A0749032/C

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

A0749032 903 bp DNA linear GSS 19-JUL-1999

HS 5575 A2 B04 T7A RPCI-11 Human Male BAC library Homo sapiens

genomic\_clone Plate=1151 Col=8 Row=I, genomic survey sequence.

A0749032.1 GI:5536190

GSS.

Homo sapiens (human)

Homo sapiens

## REFERENCE

AUTHORS

1 (bases 1 to 903)

Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holman,T.,

Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and

Hood,L.

Sequence-tagged connectors: A sequence approach to mapping and

scanning the human genome

Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)

10449764

Contact: Mahairas GG, Wallace JC, Hood L

High Throughput Sequencing Center

University of Washington

401 Queen Anne Avenue North, Seattle, WA 98109, USA

Tel: (206) 616-3618

Fax: (206) 616-3887

Email: jwallace@u.washington.edu

Clones are derived from the human BAC library RPCI-11. For BAC

library availability, please contact Pieter de Jong

(pieter@dejong.med.buffalo.edu). Clones may be purchased from

BACPAC Resources (http://bacpac.med.buffalo.edu/ordering.bac.htm)

or from Research Genetics (info@resgen.com). BAC end web server:

http://www.hsc.washington.edu

Plate: 1151 row: I column: 8

Seg primer: T7

Class: BAC ends

High quality sequence stop: 903.

Location/Qualifiers

1..903  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/clone="Plate=1151 Col=8 Row=I"  
/sex="male"  
/clone\_1ib="RPCI-11 Human Male BAC library"  
/note="Vector: pBAC3.6; Site 1: EcoRI; Site 2: EcoRI; Male blood DNA was isolated from one randomly chosen donor and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBAC3.6 vector at EcoRI sites"

## ORIGIN

Query Match 1.4%; Score 44; DB 9; Length 903;

Best Local Similarity 100.0%; Pred. No. 1.8e-09;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3073 AGATTGTCACCTGACCTCCAGCCTGGGCAACAGAGCAAGCTC 3116

Db 359 AGATTGTCACCTGACCTCCAGCCTGGGCAACAGAGCAAGCTC 316

## RESULT 192

AL570050

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AL570050 982 bp mRNA linear EST 05-APR-2004

AL570050 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

clone CS01006YK15 3-PRIME, mRNA sequence.

AL570050.3 GI:46236290

EST.

Homo sapiens (human)

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homidae; Homo.

1 (bases 1 to 982)

Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On Feb 16, 2001 this sequence version replaced gi:31291480.

Contact: Genoscope

Genoscope - Centre National de Sequencage

2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo (dt) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 6704.f

For more information about this cluster, see  
<http://www.genoscope.cns.fr/cdnas7s-CSOD1006AF08NP1ac-6704.f>.

## FEATURES

## source

1..982

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CSOD1006YK15"

/issue\_type="PLACENTA COT 25-NORMALIZED"

/clone\_1lb="Homo sapiens PLACENTA COT 25-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo (dt)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

## ORIGIN

Query Match 1.4%; Score 44; DB 1; Length 982;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3073 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGCAAGACTC 3116

351 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGCAAGACTC 394

Db

## RESULT 193

LOCUS CR622010 1527 bp mRNA linear HTC 21-JUN-2004  
 DEFINITION full-length cDNA clone CSOD1006YK15 of Placenta Cot 25-normalized  
 of Homo sapiens (human).

ACCESSION CR622010.1 GI:50502817

KEYWORDS HTC; CNSUT\_cDNA.

SOURCE Homo sapiens (human)

## ORGANISM

Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

Homidae; Homo

1 (bases 1 to 1527)

L1,W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished

Contact : Feng Liang Email : fliang@lifetech.com URL :

<http://fulllength.invitrogen.com/InvitrogenCorporation1600>

Paradey Avenue

2 (bases 1 to 1527)

Genoscope.

Direct Submission

Submitted (20-JUN-2004) Genoscope - Centre National de Sequencage :

BP 191 91006 Evry cedex - FRANCE (E-mail : segref@genoscope.cns.fr

- Web : www.genoscope.cns.fr)

1st strand cDNA was primed with a NotI-oligo (dt) primer. Five prime

end enriched, double-strand cDNA was digested with Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen.

Location/Qualifiers

1..1527

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="CSOD1006YK15"

/issue\_type="Placenta Cot 25-normalized"

/plasmid="pCMVSPORT\_6"

## ORIGIN

Query Match 1.4%; Score 44; DB 4; Length 1527;  
 Best Local Similarity 100.0%; Pred. No. 1.8e-09;  
 Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3073 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGCAAGACTC 3116

1296 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGCAAGACTC 1253

Db

## RESULT 194

## BE141837/c

## LOCUS

## DEFINITION

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## PUBMED

## COMMENT

## 108 bp mRNA linear EST 21-JUN-2000

## IL5-HT0117-170999-004-A02 HT0117 Homo sapiens CDNA, mRNA sequence.

## BE141837

## BE141837.1 GI:8604558

## EST.

## Homo sapiens

## Homo sapiens

## Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

## Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;

## Homidae; Homo

## 1 (bases 1 to 108)

## Dias Neto,E., Garcia Correa,R., Verjovski-Almeida,S., Brito,M.R.,

## Nagai,M.A., da Silva,W.J., Zago,M.A., Bordin,S., Costa,F.F.,

## Goldman,G.H., Carvalho,A.F., Matukuma,A., Bala,G.S., Simpson,D.H.,

## Brunstein,A., deOliveira,P.S., Bucher,P., Jongeneel,C.V.,

## O'Hare,M.J., Soares,F., Brentani,R.R., Reis,L.F., de Souza,S.J. and

## Simpson,A.J.

## Shotgun sequencing of the human transcriptome with ORF expressed

## sequence tags

## Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)

## 10737800

## Contact: Simpson A.J.G.

## Laboratory of Cancer Genetics

## Ludwig Institute for Cancer Research

## Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,

## Brazil

## Tel: +55-11-2704922

## Fax: +55-11-2707001

## Email: asimpson@ludwig.org.br

## This sequence was derived from the PABSP/LICR Human Cancer Genome

## Project. This entry can be seen in the following URL

<http://www.ludwig.org.br/scripts/gethtml2.pl?l1=act2-IL5-HT0117-170>

## 999-004-A02&amp;t3=1999-09-17&amp;t4=1)

## Seq primer: puc 18 forward

## High quality sequence stop: 108.

## Location/Qualifiers

## 1..108

## /organism="Homo sapiens"

## /mol\_type="mRNA"

## /db\_xref="taxon:9606"

## /def\_stage="Adult"

## /clone\_1lb="HT0117"

## /note="Organ: head, neck; Vector: puc18; Site\_1: Smal;

## Site\_2: SmaI; A mini-library was made by cloning products

## derived from ORSTS PCR (U.S. Letters Patent application

## No. 196,716 - Ludwig Institute for Cancer Research)

## profiles into the pUC 18 vector. Reverse transcription of

## tissue mRNA and cDNA amplification were performed under

## low stringency conditions."

## ORIGIN

Query Match 1.4%; Score 43; DB 2; Length 108;  
 Best Local Similarity 100.0%; Pred. No. 6e-09;  
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

3080 GCCACTGCACTCCAGCTGGGCAACAGCAAGACTCTGTCTC 3122

60 GCCACTGCACTCCAGCTGGGCAACAGCAAGACTCTGTCTC 18

Db

## RESULT 195

AA728876

LOCUS AA728876 132 bp mRNA linear EST 22-JAN-1998  
 DEFINITION nv37c08.r1 NCI CGAP Br5 Homo sapiens cDNA clone IMAGE1222382  
 similar to contains Alu repetitive element; contains element MSRI  
 repetitive element ;, mRNA sequence.

ACCESSION AA728876  
 VERSION AA728876.1 GI:2750235  
 EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Carchinhii;  
 Homnidae; Homo.  
 1 (bases 1 to 132)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished (1997)  
 CONTACT: Robert Strausberg, Ph.D.  
 Email: [cgapbs-r@mail.nih.gov](mailto:cgapbs-r@mail.nih.gov)  
 Tissue Procurement: Ilan Kirsch, M.D., Kristina A. Cole, M.D.,  
 Ph.D. student, Rodrigo F. Chuagui, M.D., Michael R. Emmert-Buck,  
 M.D., Ph.D.  
 cDNA Library Preparation: David B. Krizman, Ph.D.  
 cDNA library Arrayed by: Greg Lennon, Ph.D.  
 DNA Sequencing by: Washington University Genome Sequencing Center  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/ILNI at:  
[www.bio.lnll.gov/dbirp/image/image.html](http://www.bio.lnll.gov/dbirp/image/image.html)  
 Insert Length: 180 Std Error: 0.00  
 Seq primer: -28m13 rev1 ET from Amersham.  
 Location/Qualifiers  
 1..132  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1222382"  
 /sex="female"  
 /tissue\_type="infiltrating ductal carcinoma"  
 /lab\_host="DH10B"  
 /clone\_lib="NCI-CGAP Br5"  
 /note="Organ: breast; Vector: pAMP10; mRNA made from  
 infiltrating ductal carcinoma, cDNA made by oligo-dT  
 priming. Non-directionally cloned. Size-selected on  
 agarose gel, average insert size 600 bp."

ORIGIN  
 Query Match 1.4%; Score 43; DB 1; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-09;  
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCGACTGCACTCCAGCTTGGGCAACAGACAAAGACT 3115  
 |||||  
 Db 67 AGATTGTGCGACTGCACTCCAGCTTGGGCAACAGACAAAGACT 109

RESULT 196  
 LOCUS BG012771 132 bp mRNA linear EST 24-JAN-2001  
 DEFINITION IL5-GN0239-141200-341-d03 GN0239 Homo sapiens cDNA, mRNA sequence.  
 ACCESSION BG012771  
 VERSION BG012771.1 GI:12462298  
 EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Carchinhii;  
 Homnidae; Homo.  
 1 (bases 1 to 132)  
 Dias Neto, R., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,  
 Nagai, M.A., da Silva, M. Jr., Matsuoka, A., Bala, G.S., Simpson, D.H.,  
 Goldman, G.H., Carvalho, A.F., Matsukuma, A., Bala, G.S., Simpson, D.H.,  
 Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V.,  
 O'Hare, M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and

TITLE Simpson, A.J.  
 JOURNAL Shotgun sequencing of the human transcriptome with ORF expressed  
 PROC. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)  
 PUBLISHED 10737800  
 COMMENT Contact: Simpson A.J.G.  
 Laboratory of Cancer Genetics  
 Ludwig Institute for Cancer Research  
 Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,  
 Brazil  
 Tel: +55-11-2704922  
 Fax: +55-11-2707001  
 Email: [asimpson@ludwig.org.br](mailto:asimpson@ludwig.org.br)  
 This sequence was derived from the FAPESP/LICR Human Cancer Genome  
 Project. This entry can be seen in the following URL  
<http://www.ludwig.org.br/scripts/gethtml2.pl?cl=IL5&c2=IL5-GN0239-141200-341-d03&c3=2000-12-14&t4=1>  
 Seq primer: puc 18 forward  
 High quality sequence start: 14  
 High quality sequence stop: 132.  
 Location/Qualifiers  
 1..132  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /dev\_stage="Adult"  
 /clone\_lib="GN0239"  
 /note="Organ: placenta\_normal; Vector: puc18; Site: 1;  
 Site: 2: SmaI; A mini-library was made by cloning  
 products derived from ORESEBS PCR (U.S. Letters Patent  
 application No. 196,716 - Ludwig Institute for Cancer  
 Research) profiles into the puc 18 vector. Reverse  
 transcription of tissue mRNA and cDNA amplification were  
 performed under low stringency conditions."

ORIGIN  
 Query Match 1.4%; Score 43; DB 2; Length 132;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-09;  
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2898 GGATCACTGAGGCCAGAGTTGAGACCAAGCTTGGCCAAACT 2940  
 |||||  
 Db 127 GGATCACTGAGGCCAGAGTTGAGACCAAGCTTGGCCAAACT 85

RESULT 197  
 LOCUS CR543084 151 bp mRNA linear EST 07-JUL-2004  
 DEFINITION DKFZp470G0412.r1 470 (synonym: p1v1) Pongo pygmaeus cDNA clone  
 DKFZp470G0412.5, mRNA sequence.  
 ACCESSION CR543084  
 VERSION CR543084.1 GI:49894295  
 EST.  
 SOURCE Pongo pygmaeus (orangutan)  
 ORGANISM Pongo pygmaeus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Euarchontoglires; Primates; Carchinhii;  
 Homnidae; Pongo.  
 1 (bases 1 to 151)  
 Wambutt, R., Heubner, D., Mewes, H.W., Well, B., Auld, C., Osanger, A.,  
 Fobé, G., Han, M. and Wiemann, S.  
 Pongo pygmaeus mRNA (Wambutt, R., Heubner, D., Mewes, H.W., et al.)  
 Unpublished (2004)  
 CONTACT: MIPS  
 MIPS  
 Ingolstaedter Landstr. 1, D-85764 Neuberg, Germany  
 This is the 5' sequence of the clone insert from S. Wiemann,  
 Molecular Genome Analysis, German Cancer Research Center (DKFZ);  
 Email: [s.wiemann@dkfz-heidelberg.de](mailto:s.wiemann@dkfz-heidelberg.de); sequenced by Agowa  
 (Berlin/Germany) within the cDNA sequencing consortium of the  
 German Genome Project. This clone (DKFZp470G0412) is available at  
 the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum,  
 Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email:





## RESULT 200

AA018823/C

## LOCUS

AA018823 168 bp mRNA linear EST 30-JAN-1997  
 ze57e09.r1 Soares retina N2b4HR Homo sapiens cDNA clone  
 IMAGE:363112 5' similar to contains Alu repetitive element; mRNA  
 sequence.

## ACCESSION

AA018823

## VERSION

AA018823.1 GI:1482215

## KEYWORDS

EST.

## SOURCE

Homo sapiens (human)

## REFERENCE

AUTHORS

1 (bases 1 to 168)  
 Hillier, L., Lennon, G., Becker, M., Bonaldo, M.F., Chapell, B.,  
 Chissole, S., Dietrich, N., Dubuque, T., Favello, A., Gish, W.,  
 Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, M., Le, N.,  
 Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, L.,  
 Rohlfing, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, D.,  
 Trevaaskis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.  
 and Marra, M.

## TITLE

Generation and analysis of 280,000 human expressed sequence tags  
 Genome Res. 6 (9), 807-828 (1996)

## JOURNAL

PUBMED

## COMMENT

Contact: Wilson RK  
 Washington University School of Medicine  
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108  
 Tel: 314 286 1800  
 Fax: 314 286 1810  
 Email: est@watson.wustl.edu  
 This clone is available royalty-free through LNL; contact the  
 IMAGS Consortium (info@image.llnl.gov) for further information.  
 Insert Length: 1965 Std Error: 0.00  
 Seq primer: -28M13 rev2 from Amerham  
 High quality sequence stop: 139.

## FEATURES

source

1..168  
 location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="GDB:1279815"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:363112"  
 /sex="male"  
 /tissue\_type="retina"  
 /dev\_stage="55 year old"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_1lb="Soares retina N2b4HR"  
 /note="Organ: eye; Vector: pT73D (Pharmacia) with a  
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer [5'  
 TGTTCACATCTGAAGTGGAGCGCGCGCTTTTCTTTTCTTTT 3'],  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified pT73 vector  
 (Pharmacia). The retinas were obtained from a 55 year old  
 Caucasian and local cellular poly(A)+ RNA was extracted 6  
 hrs after their removal. The retina RNA was kindly  
 provided by Roderick R. McInnes M.D. Ph.D. from the  
 University of Toronto. Library constructed by Bento  
 Soares and M. Fatima Bonaldo."

## ORIGIN

## Query Match

Best Local Similarity 100.0%; Score 43; DB 1; Length 168;  
 Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3080 GCCACTGCCTCAGCGCTGGCAACAGACAAAGACTCTGTCTC 3122

DB 128 GCCACTGCCTCAGCGCTGGCAACAGACAAAGACTCTGTCTC 86



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2006, 05:27:55 / Search time 531 Seconds  
(without alignments)  
10451.125 Million cell updates/sec

Title: US-10-760-320A-102  
Perfect score: 3122  
Sequence: 1 actagaggtggtggttagcgc.....acagagaactctgtctc 3122

Scoring table: **OLIGO\_NUC**  
Gapop 60.0, Gapext 60.0

Searched: 1303057 seqs, 888780828 residues

Word size: 1

Total number of hits satisfying chosen parameters: 2605320

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Listing first 500 summaries

Database: Issued Patents NA:  
1: /cgn2\_6/prodata/1/ina/1/COMB.seq:\*  
2: /cgn2\_6/prodata/1/ina/5/COMB.seq:\*  
3: /cgn2\_6/prodata/1/ina/6A/COMB.seq:\*  
4: /cgn2\_6/prodata/1/ina/6B/COMB.seq:\*  
5: /cgn2\_6/prodata/1/ina/H/COMB.seq:\*  
6: /cgn2\_6/prodata/1/ina/PCITUS/COMB.seq:\*  
7: /cgn2\_6/prodata/1/ina/PP/COMB.seq:\*  
8: /cgn2\_6/prodata/1/ina/RE/COMB.seq:\*  
9: /cgn2\_6/prodata/1/ina/backfilest1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	52	1.7	381	3	US-09-513-999C-3793
C 2	52	1.7	601	3	US-09-949-016-72610
C 3	52	1.7	601	3	US-09-949-016-75710
C 4	52	1.7	601	3	US-09-949-016-75723
C 5	52	1.7	15415	3	US-09-949-016-13931
C 6	52	1.7	15415	3	US-09-949-016-13931
C 7	52	1.7	25869	3	US-09-949-016-13932
C 8	52	1.7	77586	3	US-09-949-016-13920
C 9	52	1.7	77586	3	US-09-949-016-13920
C 10	52	1.7	92139	3	US-09-918-686-1
C 11	52	1.7	120727	3	US-09-949-016-15787
C 12	52	1.7	120727	3	US-09-949-016-15788
C 13	52	1.7	155266	3	US-09-949-016-15870
C 14	51	1.6	35688	3	US-09-949-016-16873
C 15	51	1.6	156942	3	US-09-949-016-12227
C 16	51	1.6	255	3	US-09-949-016-15946
C 17	50	1.6	288	3	US-09-973-278-870
C 18	50	1.6	601	3	US-09-949-016-47563
C 19	50	1.6	601	3	US-09-949-016-125579
C 20	50	1.6	601	3	US-09-949-016-125579
C 21	50	1.6	601	3	US-09-949-016-125580
C 22	50	1.6	601	3	US-09-949-016-125581
C 23	50	1.6	601	3	US-09-949-016-163881
C 24	50	1.6	8705	3	US-09-544-398B-10

C 25	50	1.6	8705	3	US-09-543-771B-10	Sequence 10, App1
C 26	50	1.6	36954	3	US-09-949-016-15304	Sequence 15304, A
C 27	50	1.6	39154	3	US-09-949-016-13384	Sequence 12384, A
C 28	50	1.6	39154	3	US-09-949-016-12801	Sequence 12801, A
C 29	50	1.6	39443	3	US-09-949-016-14326	Sequence 14326, A
C 30	50	1.6	39443	3	US-09-949-016-14327	Sequence 14327, A
C 31	50	1.6	76401	3	US-09-949-016-17153	Sequence 17153, A
C 32	50	1.6	87190	3	US-09-949-016-16335	Sequence 16335, A
C 33	50	1.6	95648	3	US-09-949-016-11319	Sequence 11319, A
C 34	50	1.6	140844	3	US-09-949-016-14199	Sequence 14199, A
C 35	50	1.6	197336	3	US-09-949-016-12881	Sequence 12881, A
C 36	50	1.6	197337	3	US-09-949-016-14376	Sequence 14376, A
C 37	50	1.6	197337	3	US-09-949-002-738	Sequence 738, App
C 38	50	1.6	234287	3	US-09-949-002-687	Sequence 687, App
C 39	50	1.6	234288	3	US-09-949-016-17272	Sequence 17272, A
C 40	50	1.6	234288	3	US-09-949-002-841	Sequence 841, App
C 41	50	1.6	254964	3	US-09-949-016-12583	Sequence 12583, A
C 42	50	1.6	254964	3	US-09-949-016-17392	Sequence 17392, A
C 43	47	1.5	601	3	US-09-949-016-16469	Sequence 16469, A
C 44	47	1.5	12880	3	US-09-949-016-16733	Sequence 16733, A
C 45	47	1.5	45833	3	US-09-949-016-14330	Sequence 14330, A
C 46	47	1.5	45833	3	US-09-949-016-14331	Sequence 14331, A
C 47	47	1.5	45833	3	US-09-949-016-14332	Sequence 14332, A
C 48	47	1.5	45833	3	US-09-949-016-14333	Sequence 14333, A
C 49	47	1.5	67386	3	US-09-949-016-16519	Sequence 16519, A
C 50	47	1.5	108310	3	US-09-949-016-16366	Sequence 16366, A
C 51	47	1.5	139849	3	US-09-949-016-17030	Sequence 17030, A
C 52	46	1.5	601	3	US-09-949-016-65112	Sequence 65112, A
C 53	46	1.5	601	3	US-09-949-016-65113	Sequence 65113, A
C 54	46	1.5	601	3	US-09-949-016-65114	Sequence 65114, A
C 55	46	1.5	601	3	US-09-949-016-137514	Sequence 137514, A
C 56	46	1.5	601	3	US-09-949-016-15725	Sequence 15725, A
C 57	46	1.5	601	3	US-09-949-016-15726	Sequence 15726, A
C 58	46	1.5	601	3	US-09-949-016-169304	Sequence 169304, A
C 59	46	1.5	601	3	US-09-949-016-169305	Sequence 169305, A
C 60	46	1.5	601	3	US-09-949-016-169306	Sequence 169306, A
C 61	46	1.5	601	3	US-09-949-016-198268	Sequence 198268, A
C 62	46	1.5	601	3	US-09-949-016-198269	Sequence 198269, A
C 63	46	1.5	601	3	US-09-949-016-198270	Sequence 198270, A
C 64	46	1.5	601	3	US-09-949-016-206260	Sequence 206260, A
C 65	46	1.5	601	3	US-09-949-016-206261	Sequence 206261, A
C 66	46	1.5	601	3	US-09-949-016-206262	Sequence 206262, A
C 67	46	1.5	177004	3	US-09-949-016-14878	Sequence 14878, A
C 68	46	1.5	27600	3	US-09-949-016-15290	Sequence 15290, A
C 69	46	1.5	29860	3	US-09-949-016-11978	Sequence 11978, A
C 70	46	1.5	30054	3	US-09-949-016-16100	Sequence 16100, A
C 71	46	1.5	40905	3	US-09-949-016-16864	Sequence 16864, A
C 72	46	1.5	60917	3	US-09-949-016-16057	Sequence 16057, A
C 73	46	1.5	60917	3	US-09-949-016-16057	Sequence 16057, A
C 74	46	1.5	96739	3	US-09-949-016-14736	Sequence 14736, A
C 75	46	1.5	96739	3	US-09-949-016-15606	Sequence 15606, A
C 76	46	1.5	116955	3	US-09-949-016-17565	Sequence 17565, A
C 77	46	1.5	118136	3	US-09-949-016-14339	Sequence 14339, A
C 78	46	1.5	144322	3	US-09-949-016-15316	Sequence 15316, A
C 79	46	1.5	177293	3	US-09-949-016-16513	Sequence 16513, A
C 80	46	1.5	240157	3	US-09-949-016-16264	Sequence 16264, A
C 81	46	1.5	253345	3	US-09-949-016-16566	Sequence 16566, A
C 82	46	1.5	253345	3	US-09-949-016-16566	Sequence 16566, A
C 83	46	1.5	331814	3	US-09-949-016-12639	Sequence 12639, A
C 84	46	1.5	331814	3	US-09-949-016-12639	Sequence 12639, A
C 85	46	1.5	331814	3	US-09-949-016-12639	Sequence 12639, A
C 86	46	1.5	767677	3	US-09-949-016-13147	Sequence 13147, A
C 87	46	1.5	767677	3	US-09-949-016-13147	Sequence 13147, A
C 88	45	1.4	478	3	US-09-621-976-15071	Sequence 15071, A
C 89	45	1.4	601	3	US-09-949-016-30237	Sequence 30237, A
C 90	45	1.4	601	3	US-09-949-016-30238	Sequence 30238, A
C 91	45	1.4	601	3	US-09-949-016-30239	Sequence 30239, A
C 92	45	1.4	601	3	US-09-949-016-65461	Sequence 65461, A
C 93	45	1.4	601	3	US-09-949-016-72557	Sequence 72557, A
C 94	45	1.4	601	3	US-09-949-016-82095	Sequence 82095, A
C 95	45	1.4	601	3	US-09-949-016-82096	Sequence 82096, A
C 96	45	1.4	601	3	US-09-949-016-82097	Sequence 82097, A
C 97	45	1.4	601	3	US-09-949-016-145830	Sequence 145830, A

C 98	-45	1.4	601	3	US-09-949-016-146098	Sequence 146098,	C 111	44	1.4	601	3	US-09-949-016-121487
C 99	45	1.4	601	3	US-09-949-016-146366	Sequence 146366,	C 112	44	1.4	601	3	US-09-949-016-121488
C 100	45	1.4	601	3	US-09-949-016-146442	Sequence 146442,	C 113	44	1.4	601	3	US-09-949-016-133933
C 101	45	1.4	601	3	US-09-949-016-159130	Sequence 169130,	C 114	44	1.4	601	3	US-09-949-016-133934
C 102	45	1.4	601	3	US-09-949-016-159131	Sequence 169131,	C 115	44	1.4	601	3	US-09-949-016-143264
C 103	45	1.4	601	3	US-09-949-016-203717	Sequence 203717,	C 116	44	1.4	601	3	US-09-949-016-143265
C 104	45	1.4	601	3	US-09-949-016-203718	Sequence 203718,	C 117	44	1.4	601	3	US-09-949-016-144443
C 105	45	1.4	601	3	US-09-949-016-203719	Sequence 203719,	C 118	44	1.4	601	3	US-09-949-016-144444
C 106	45	1.4	601	3	US-09-949-002-4958	Sequence 4958, Ap	C 119	44	1.4	601	3	US-09-949-016-150959
C 107	45	1.4	601	3	US-09-949-002-4959	Sequence 4959, Ap	C 120	44	1.4	601	3	US-09-949-016-163006
C 108	45	1.4	601	3	US-09-949-002-9370	Sequence 9370, Ap	C 121	44	1.4	601	3	US-09-949-016-165548
C 109	45	1.4	8429	3	US-09-949-016-14483	Sequence 14483, A	C 122	44	1.4	601	3	US-09-949-016-199110
C 110	45	1.4	9123	3	US-09-949-016-14852	Sequence 14852, A	C 123	44	1.4	601	3	US-09-949-016-202729
C 111	45	1.4	9176	3	US-09-949-016-17432	Sequence 17432, A	C 124	44	1.4	601	3	Sequence 205472,
C 112	45	1.4	10597	3	US-09-949-016-13766	Sequence 13766, A	C 125	44	1.4	601	3	Sequence 204472,
C 113	45	1.4	15975	3	US-09-949-016-17469	Sequence 17469, A	C 126	44	1.4	601	3	Sequence 10588, A
C 114	45	1.4	16595	3	US-09-146-053-7	Sequence 7, App1.1	C 127	44	1.4	956	3	Sequence 56, App1
C 115	45	1.4	19296	3	US-09-949-016-16393	Sequence 16393, A	C 128	44	1.4	7480	3	Sequence 56, App1
C 116	45	1.4	19943	3	US-09-949-016-15920	Sequence 15920, A	C 129	44	1.4	7567	3	Sequence 11999, A
C 117	45	1.4	19945	3	US-09-949-016-12553	Sequence 12553, A	C 130	44	1.4	7720	3	Sequence 16486, A
C 118	45	1.4	27845	3	US-09-949-016-13069	Sequence 13069, A	C 131	44	1.4	7720	3	Sequence 5, App1.1
C 119	45	1.4	33099	3	US-09-949-016-16094	Sequence 16094, A	C 132	44	1.4	7720	3	Sequence 17443, A
C 120	45	1.4	36223	3	US-09-949-016-14417	Sequence 14417, A	C 133	44	1.4	10720	3	Sequence 12354, A
C 121	45	1.4	36546	3	US-09-949-016-17543	Sequence 17543, A	C 134	44	1.4	15116	3	Sequence 16260, A
C 122	45	1.4	40090	3	US-09-820-004-3	Sequence 3, App1.1	C 135	44	1.4	15117	3	Sequence 592, App
C 123	45	1.4	40147	3	US-09-949-016-13374	Sequence 12374, A	C 136	44	1.4	16851	3	Sequence 766, App
C 124	45	1.4	40379	3	US-09-949-016-14128	Sequence 14128, A	C 137	44	1.4	16891	3	Sequence 16523, A
C 125	45	1.4	40877	3	US-09-949-016-14128	Sequence 17495, A	C 138	44	1.4	18891	3	Sequence 15113, A
C 126	45	1.4	40908	3	US-09-949-002-845	Sequence 845, App	C 139	44	1.4	19384	3	Sequence 15677, A
C 127	45	1.4	40944	3	US-09-949-002-603	Sequence 603, App	C 140	44	1.4	19389	3	Sequence 651, App
C 128	45	1.4	47781	3	US-09-949-016-15492	Sequence 16492, A	C 141	44	1.4	20674	3	Sequence 651, App
C 129	45	1.4	47781	3	US-09-949-016-15492	Sequence 16492, A	C 142	44	1.4	20674	3	Sequence 16471, A
C 130	45	1.4	47781	3	US-09-949-016-15494	Sequence 16494, A	C 143	44	1.4	22906	3	Sequence 12691, A
C 131	45	1.4	49677	3	US-09-949-002-817	Sequence 817, App	C 144	44	1.4	23160	3	Sequence 13651, A
C 132	45	1.4	64984	3	US-09-949-016-15254	Sequence 15254, A	C 145	44	1.4	25162	3	Sequence 13066, A
C 133	45	1.4	66627	3	US-09-949-016-12630	Sequence 12630, A	C 146	44	1.4	26072	3	Sequence 17222, A
C 134	45	1.4	66628	3	US-09-949-016-16112	Sequence 16112, A	C 147	44	1.4	26072	3	Sequence 12825, A
C 135	45	1.4	70313	3	US-09-949-002-714	Sequence 714, App	C 148	44	1.4	28809	3	Sequence 15326, A
C 136	45	1.4	72347	3	US-09-949-002-633	Sequence 633, App	C 149	44	1.4	29930	3	Sequence 17532, A
C 137	45	1.4	80858	3	US-09-949-016-12859	Sequence 12859, A	C 150	44	1.4	33444	3	Sequence 16306, A
C 138	45	1.4	80859	3	US-09-949-016-15715	Sequence 15715, A	C 151	44	1.4	33870	3	Sequence 12518, A
C 139	45	1.4	87648	3	US-09-949-016-13655	Sequence 13655, A	C 152	44	1.4	34315	3	Sequence 12518, A
C 140	45	1.4	91379	3	US-09-949-016-15146	Sequence 15146, A	C 153	44	1.4	34315	3	Sequence 12561, A
C 141	45	1.4	93532	3	US-09-949-016-15944	Sequence 15944, A	C 154	44	1.4	34315	3	Sequence 12561, A
C 142	45	1.4	94142	3	US-09-949-016-15553	Sequence 16553, A	C 155	44	1.4	38503	3	Sequence 14051, A
C 143	45	1.4	96690	3	US-09-949-016-17103	Sequence 17103, A	C 156	44	1.4	37315	3	Sequence 14254, A
C 144	45	1.4	103750	3	US-09-949-016-13119	Sequence 13119, A	C 157	44	1.4	45986	3	Sequence 16716, A
C 145	45	1.4	112465	3	US-09-949-016-16424	Sequence 16424, A	C 158	44	1.4	46961	3	Sequence 16716, A
C 146	45	1.4	120609	3	US-09-949-016-13915	Sequence 13915, A	C 159	44	1.4	56782	3	Sequence 16551, A
C 147	45	1.4	140725	3	US-09-949-016-10704	Sequence 17074, A	C 160	44	1.4	60137	3	Sequence 14735, A
C 148	45	1.4	155266	3	US-09-949-016-13870	Sequence 13870, A	C 161	44	1.4	60137	3	Sequence 14912, A
C 149	45	1.4	162841	3	US-09-949-016-13733	Sequence 13733, A	C 162	44	1.4	64291	3	Sequence 12604, A
C 150	45	1.4	174639	3	US-09-949-016-16509	Sequence 16509, A	C 163	44	1.4	64291	3	Sequence 12604, A
C 151	45	1.4	205044	3	US-09-949-016-15851	Sequence 15851, A	C 164	44	1.4	65848	3	Sequence 13885, A
C 152	45	1.4	205044	3	US-09-949-016-15852	Sequence 15852, A	C 165	44	1.4	65848	3	Sequence 16406, A
C 153	45	1.4	205044	3	US-09-949-016-15853	Sequence 15853, A	C 166	44	1.4	68768	3	Sequence 15532, A
C 154	45	1.4	218940	3	US-09-949-016-17539	Sequence 17539, A	C 167	44	1.4	82000	3	Sequence 15532, A
C 155	45	1.4	223471	3	US-09-949-016-1287	Sequence 12387, A	C 168	44	1.4	82000	3	Sequence 13424, A
C 156	45	1.4	223471	3	US-09-949-016-12725	Sequence 12725, A	C 169	44	1.4	85850	3	Sequence 12017, A
C 157	45	1.4	223471	3	US-09-949-016-12725	Sequence 12725, A	C 170	44	1.4	85850	3	Sequence 12017, A
C 158	45	1.4	223471	3	US-09-949-016-12725	Sequence 12725, A	C 171	44	1.4	85850	3	Sequence 12017, A
C 159	45	1.4	330	3	US-09-513-999C-25910	Sequence 25910, A	C 172	44	1.4	85850	3	Sequence 12017, A
C 160	45	1.4	330	3	US-09-513-999C-28305	Sequence 28305, A	C 173	44	1.4	85850	3	Sequence 12017, A
C 161	45	1.4	601	3	US-09-949-016-32113	Sequence 32113, A	C 174	44	1.4	85850	3	Sequence 12017, A
C 162	45	1.4	601	3	US-09-949-016-36581	Sequence 36581, A	C 175	44	1.4	85850	3	Sequence 12017, A
C 163	45	1.4	601	3	US-09-949-016-38281	Sequence 38281, A	C 176	44	1.4	85850	3	Sequence 12017, A
C 164	45	1.4	601	3	US-09-949-016-46100	Sequence 46100, A	C 177	44	1.4	85850	3	Sequence 12017, A
C 165	45	1.4	601	3	US-09-949-016-46119	Sequence 46119, A	C 178	44	1.4	85850	3	Sequence 12017, A
C 166	45	1.4	601	3	US-09-949-016-71064	Sequence 71064, A	C 179	44	1.4	85850	3	Sequence 12017, A
C 167	45	1.4	601	3	US-09-949-016-71065	Sequence 71065, A	C 180	44	1.4	85850	3	Sequence 12017, A
C 168	45	1.4	601	3	US-09-949-016-71066	Sequence 71066, A	C 181	44	1.4	85850	3	Sequence 12017, A
C 169	45	1.4	601	3	US-09-949-016-71067	Sequence 71067, A	C 182	44	1.4	85850	3	Sequence 12017, A
C 170	45	1.4	601	3	US-09-949-016-71068	Sequence 71068, A	C 183	44	1.4	85850	3	Sequence 12017, A
							C 184	44	1.4	85850	3	Sequence 12017, A
							C 185	44	1.4	85850	3	Sequence 12017, A
							C 186	44	1.4	85850	3	Sequence 12017, A
							C 187	44	1.4	85850	3	Sequence 12017, A
							C 188	44	1.4	85850	3	Sequence 12017, A
							C 189	44	1.4	85850	3	Sequence 12017, A
							C 190	44	1.4	85850	3	Sequence 12017, A
							C 191	44	1.4	85850	3	Sequence 12017, A
							C 192	44	1.4	85850	3	Sequence 12017, A
							C 193	44	1.4	85850	3	Sequence 12017, A
							C 194	44	1.4	85850	3	Sequence 12017, A
							C 195	44	1.4	85850	3	Sequence 12017, A
							C 196	44	1.4	85850	3	Sequence 12017, A
							C 197	44	1.4	85850	3	Sequence 12017, A
							C 198	44	1.4	85850	3	Sequence 12017, A
							C 199	44	1.4	85850	3	Sequence 12017, A
							C 200	44	1.4	85850	3	Sequence 12017, A
							C 201	44	1.4	85850	3	Sequence 12017, A
							C 202	44	1.4	85850	3	Sequence 12017, A
							C 203	44	1.4	85850	3	Sequence 12017, A
							C 204	44	1.4	85850	3	Sequence 12017, A
							C 205	44	1.4	85850	3	Sequence 12017, A
							C 206	44	1.4	85850	3	Sequence 12017, A
							C 207	44	1.4	85850	3	Sequence 12017, A
							C 208	44	1.4	85850	3	Sequence 12017, A
							C 209	44	1.4	85850	3	Sequence 12017, A
							C 210	44	1.4	85850	3	Sequence 12017, A
							C 21					

C 244	44	1.4	114183	3	US-09-949-002-849	Sequence 849, App	C 317	43	1.4	365	3	US-09-513-999C-21034	Sequence 21034, A
C 245	44	1.4	117410	3	US-09-949-016-12262	Sequence 12262, A	318	43	1.4	601	3	US-09-949-016-17174	Sequence 17714, A
C 246	44	1.4	120727	3	US-09-949-016-15787	Sequence 15787, A	319	43	1.4	601	3	US-09-949-016-20309	Sequence 20310, A
C 247	44	1.4	120727	3	US-09-949-016-15788	Sequence 15788, A	320	43	1.4	601	3	US-09-949-016-20310	Sequence 20310, A
C 248	44	1.4	127546	3	US-09-949-002-624	Sequence 624, App	321	43	1.4	601	3	US-09-949-016-21707	Sequence 21707, A
C 249	44	1.4	133613	3	US-09-949-016-15824	Sequence 15824, A	322	43	1.4	601	3	US-09-949-016-21708	Sequence 21708, A
C 250	44	1.4	134408	3	US-09-949-016-13841	Sequence 13841, A	323	43	1.4	601	3	US-09-949-016-21709	Sequence 21709, A
C 251	44	1.4	134434	3	US-09-949-016-17362	Sequence 17362, A	324	43	1.4	601	3	US-09-949-016-26397	Sequence 26397, A
C 252	44	1.4	139257	3	US-09-920-671-11	Sequence 11, App1	325	43	1.4	601	3	US-09-949-016-35107	Sequence 35107, A
C 253	44	1.4	139936	3	US-09-949-016-11782	Sequence 11782, A	326	43	1.4	601	3	US-09-949-016-35108	Sequence 35108, A
C 254	44	1.4	139952	3	US-09-949-016-13280	Sequence 13280, A	327	43	1.4	601	3	US-09-949-016-42327	Sequence 45237, A
C 255	44	1.4	146307	3	US-09-949-016-14881	Sequence 14881, A	328	43	1.4	601	3	US-09-949-016-46203	Sequence 46203, A
C 256	44	1.4	146307	3	US-09-949-016-14882	Sequence 14882, A	329	43	1.4	601	3	US-09-949-016-55368	Sequence 55368, A
C 257	44	1.4	146307	3	US-09-949-016-14883	Sequence 14883, A	330	43	1.4	601	3	US-09-949-016-55369	Sequence 55369, A
C 258	44	1.4	146307	3	US-09-949-016-14884	Sequence 14884, A	331	43	1.4	601	3	US-09-949-016-66550	Sequence 66550, A
C 259	44	1.4	146307	3	US-09-949-016-14885	Sequence 14885, A	332	43	1.4	601	3	US-09-949-016-89155	Sequence 89155, A
C 260	44	1.4	146307	3	US-09-949-016-14886	Sequence 14886, A	333	43	1.4	601	3	US-09-949-016-103330	Sequence 103330, A
C 261	44	1.4	146307	3	US-09-949-016-14887	Sequence 14887, A	334	43	1.4	601	3	US-09-949-016-103331	Sequence 103331, A
C 262	44	1.4	146307	3	US-09-949-016-14888	Sequence 14888, A	335	43	1.4	601	3	US-09-949-016-105600	Sequence 105600, A
C 263	44	1.4	148405	3	US-09-949-016-11747	Sequence 11747, A	336	43	1.4	601	3	US-09-949-016-132803	Sequence 132803, A
C 264	44	1.4	148405	3	US-09-949-016-12835	Sequence 12835, A	337	43	1.4	601	3	US-09-949-016-132804	Sequence 132804, A
C 265	44	1.4	148405	3	US-09-949-016-12836	Sequence 12836, A	338	43	1.4	601	3	US-09-949-016-137031	Sequence 137031, A
C 266	44	1.4	148405	3	US-09-949-016-12837	Sequence 12837, A	339	43	1.4	601	3	US-09-949-016-137032	Sequence 137032, A
C 267	44	1.4	152481	3	US-09-949-016-12521	Sequence 12521, A	340	43	1.4	601	3	US-09-949-016-138675	Sequence 138675, A
C 268	44	1.4	152481	3	US-09-949-016-12775	Sequence 12775, A	341	43	1.4	601	3	US-09-949-016-142312	Sequence 142312, A
C 269	44	1.4	152822	3	US-09-949-016-17518	Sequence 17518, A	342	43	1.4	601	3	US-09-949-016-142313	Sequence 142313, A
C 270	44	1.4	152822	3	US-09-949-016-17519	Sequence 17519, A	343	43	1.4	601	3	US-09-949-016-14926	Sequence 14926, A
C 271	44	1.4	157822	3	US-09-949-016-16723	Sequence 16723, A	344	43	1.4	601	3	US-09-949-016-167908	Sequence 167908, A
C 272	44	1.4	205044	3	US-09-949-016-15851	Sequence 15851, A	345	43	1.4	601	3	US-09-949-016-167909	Sequence 167909, A
C 273	44	1.4	205044	3	US-09-949-016-15852	Sequence 15852, A	346	43	1.4	601	3	US-09-949-016-175295	Sequence 175295, A
C 274	44	1.4	205044	3	US-09-949-016-15853	Sequence 15853, A	347	43	1.4	601	3	US-09-949-016-179586	Sequence 179586, A
C 275	44	1.4	211049	3	US-09-949-016-15770	Sequence 15770, A	348	43	1.4	601	3	US-09-949-016-179587	Sequence 179587, A
C 276	44	1.4	221958	3	US-09-949-016-13173	Sequence 13173, A	349	43	1.4	601	3	US-09-949-016-179588	Sequence 179588, A
C 277	44	1.4	221966	3	US-09-949-016-15498	Sequence 15498, A	350	43	1.4	601	3	US-09-949-016-183547	Sequence 183547, A
C 278	44	1.4	223471	3	US-09-949-016-13387	Sequence 12387, A	351	43	1.4	601	3	US-09-949-016-169642	Sequence 169642, A
C 279	44	1.4	223471	3	US-09-949-016-12724	Sequence 12724, A	352	43	1.4	5232	3	US-09-949-016-17173	Sequence 3, App1
C 280	44	1.4	223471	3	US-09-949-016-12725	Sequence 12725, A	353	43	1.4	5232	3	US-08-800-929A-3	Sequence 3, App1
C 281	44	1.4	237510	3	US-09-949-016-14273	Sequence 14273, A	354	43	1.4	5232	3	US-09-617-053A-3	Sequence 3, App1
C 282	44	1.4	239527	3	US-09-949-016-15980	Sequence 15980, A	355	43	1.4	5372	3	US-09-949-016-14484	Sequence 14484, A
C 283	44	1.4	250352	3	US-09-949-016-14724	Sequence 14724, A	356	43	1.4	5386	3	US-09-949-016-11885	Sequence 11885, A
C 284	44	1.4	275110	3	US-09-949-016-16706	Sequence 12706, A	357	43	1.4	5749	3	US-09-949-016-15441	Sequence 15441, A
C 285	44	1.4	275110	3	US-09-949-016-16707	Sequence 16707, A	358	43	1.4	5749	3	US-09-949-016-15442	Sequence 15442, A
C 286	44	1.4	304533	3	US-09-949-016-15371	Sequence 15371, A	359	43	1.4	10013	3	US-09-949-016-16474	Sequence 16474, A
C 287	44	1.4	304533	3	US-09-949-016-15372	Sequence 15372, A	360	43	1.4	15033	3	US-09-949-016-15916	Sequence 15916, A
C 288	44	1.4	385136	3	US-09-949-016-16073	Sequence 16073, A	361	43	1.4	20908	3	US-09-949-016-12684	Sequence 12684, A
C 289	44	1.4	393753	3	US-09-949-016-14573	Sequence 14573, A	362	43	1.4	20908	3	US-09-949-016-12684	Sequence 12684, A
C 290	44	1.4	393753	3	US-09-949-016-14574	Sequence 14574, A	363	43	1.4	22946	3	US-09-949-016-13381	Sequence 13381, A
C 291	44	1.4	784019	3	US-09-949-016-14033	Sequence 14033, A	364	43	1.4	25370	3	US-09-949-016-12109	Sequence 12109, A
C 292	44	1.4	818128	3	US-09-949-016-14546	Sequence 14546, A	365	43	1.4	33318	3	US-09-949-016-13798	Sequence 13798, A
C 293	44	1.4	818128	3	US-09-949-016-14547	Sequence 14547, A	366	43	1.4	31229	3	US-09-949-016-12619	Sequence 12619, A
C 294	44	1.4	818128	3	US-09-949-016-14548	Sequence 14548, A	367	43	1.4	31231	3	US-09-949-016-15766	Sequence 15766, A
C 295	44	1.4	818128	3	US-09-949-016-14549	Sequence 14549, A	368	43	1.4	32584	3	US-09-949-016-16766	Sequence 16766, A
C 296	44	1.4	818128	3	US-09-949-016-14550	Sequence 14550, A	369	43	1.4	33152	3	US-09-949-016-16719	Sequence 16719, A
C 297	44	1.4	818128	3	US-09-949-016-14551	Sequence 14551, A	370	43	1.4	38161	3	US-09-949-016-16218	Sequence 16218, A
C 298	44	1.4	818128	3	US-09-949-016-14552	Sequence 14552, A	371	43	1.4	33318	3	US-09-949-016-13798	Sequence 13798, A
C 299	44	1.4	818128	3	US-09-949-016-14553	Sequence 14553, A	372	43	1.4	42246	3	US-09-949-016-17008	Sequence 17008, A
C 300	44	1.4	818128	3	US-09-949-016-14554	Sequence 14554, A	373	43	1.4	44527	3	US-09-949-016-17525	Sequence 17525, A
C 301	44	1.4	818128	3	US-09-949-016-14555	Sequence 14555, A	374	43	1.4	45427	3	US-09-949-016-16243	Sequence 16243, A
C 302	44	1.4	818128	3	US-09-949-016-14556	Sequence 14556, A	375	43	1.4	48763	3	US-09-916-204-3	Sequence 3, App1
C 303	44	1.4	818128	3	US-09-949-016-14557	Sequence 14557, A	376	43	1.4	48763	3	US-10-282-048-3	Sequence 3, App1
C 304	44	1.4	818128	3	US-09-949-016-14558	Sequence 14558, A	377	43	1.4	49309	3	US-09-949-016-12680	Sequence 12680, A
C 305	44	1.4	818128	3	US-09-949-016-14559	Sequence 14559, A	378	43	1.4	50309	3	US-09-949-016-13084	Sequence 13084, A
C 306	44	1.4	818128	3	US-09-949-016-14560	Sequence 14560, A	379	43	1.4	45735	3	US-09-949-016-12858	Sequence 12858, A
C 307	44	1.4	818128	3	US-09-949-016-14561	Sequence 14561, A	380	43	1.4	50776	3	US-09-949-016-15438	Sequence 15438, A
C 308	44	1.4	818128	3	US-09-949-016-14562	Sequence 14562, A	381	43	1.4	51508	3	US-09-949-016-16681	Sequence 16681, A
C 309	44	1.4	818128	3	US-09-949-016-14564	Sequence 14564, A	382	43	1.4	57936	3	US-09-949-016-11935	Sequence 11935, A
C 310	44	1.4	818128	3	US-09-949-016-14565	Sequence 14565, A	383	43	1.4	57936	3	US-09-949-016-16921	Sequence 16921, A
C 311	44	1.4	818128	3	US-09-949-016-14566	Sequence 14566, A	384	43	1.4	58014	3	US-09-949-016-17448	Sequence 17448, A
C 312	44	1.4	818128	3	US-09-949-016-14567	Sequence 14567, A	385	43	1.4	59123	3	US-09-949-016-12177	Sequence 12177, A
C 313	44	1.4	828152	3	US-09-949-016-12777	Sequence 12777, A	386	43	1.4	59978	3	US-09-949-016-12259	Sequence 12259, A
C 314	43	1.4	117	3	US-09-513-999C-17724	Sequence 17724, A	387	43	1.4	59978	3	US-09-949-016-15397	Sequence 15397, A
C 315	43	1.4	354	3	US-09-973-278-811	Sequence 811, App	388	43	1.4	63386	3	US-09-949-016-12823	Sequence 12823, A
C 316	43	1.4	356	3	US-09-973-278-804	Sequence 804, App	389	43	1.4	75378	3	US-09-949-016-17140	Sequence 17140, A

C 390	43	1.4	76118	3	US-09-949-016-15593	Sequence 15593, A
C 391	43	1.4	76269	3	US-09-949-016-14603	Sequence 14603, A
C 392	43	1.4	76553	3	US-09-949-016-14332	Sequence 13432, A
C 393	43	1.4	83210	3	US-09-949-016-14209	Sequence 14209, A
C 394	43	1.4	87752	3	US-09-949-016-16807	Sequence 16807, A
C 395	43	1.4	90618	3	US-09-949-016-15964	Sequence 15964, A
C 396	43	1.4	91559	3	US-09-949-016-13581	Sequence 12581, A
C 397	43	1.4	91559	3	US-09-949-016-13701	Sequence 13701, A
C 398	43	1.4	93364	3	US-09-949-016-1890	Sequence 14890, A
C 399	43	1.4	93894	3	US-09-949-016-15629	Sequence 13629, A
C 400	43	1.4	99797	3	US-09-949-016-15255	Sequence 15255, A
C 401	43	1.4	106256	3	US-09-949-016-16858	Sequence 16858, A
C 402	43	1.4	109250	3	US-09-949-016-12530	Sequence 12530, A
C 403	43	1.4	109251	3	US-09-949-016-17321	Sequence 17321, A
C 404	43	1.4	113379	3	US-09-949-016-17561	Sequence 17561, A
C 405	43	1.4	113379	3	US-09-949-016-17561	Sequence 17561, A
C 406	43	1.4	113379	3	US-09-949-016-17562	Sequence 17562, A
C 407	43	1.4	113379	3	US-09-949-016-17562	Sequence 17562, A
C 408	43	1.4	120213	3	US-09-949-016-13304	Sequence 13304, A
C 409	43	1.4	120217	3	US-09-949-016-13260	Sequence 12260, A
C 410	43	1.4	122772	3	US-09-949-016-14132	Sequence 14132, A
C 411	43	1.4	129899	3	US-09-949-016-14684	Sequence 14684, A
C 412	43	1.4	144596	3	US-09-949-016-11749	Sequence 11749, A
C 413	43	1.4	144596	3	US-09-949-016-13035	Sequence 13035, A
C 414	43	1.4	148156	3	US-09-949-016-11776	Sequence 11776, A
C 415	43	1.4	154626	3	US-09-949-016-14000	Sequence 14000, A
C 416	43	1.4	160018	3	US-09-949-016-12617	Sequence 12617, A
C 417	43	1.4	160018	3	US-09-949-016-15994	Sequence 15994, A
C 418	43	1.4	160552	3	US-09-593-828-11	Sequence 11, App1
C 419	43	1.4	173992	3	US-09-949-016-13379	Sequence 13379, A
C 420	43	1.4	174493	3	US-09-804-471A-3	Sequence 3, App1
C 421	43	1.4	174493	3	US-10-238-709-3	Sequence 3, App1
C 422	43	1.4	174493	3	US-10-724-594-3	Sequence 3, App1
C 423	43	1.4	17751	3	US-09-949-016-15841	Sequence 15841, A
C 424	43	1.4	194714	3	US-09-949-016-11869	Sequence 11869, A
C 425	43	1.4	194889	3	US-09-949-016-15654	Sequence 15654, A
C 426	43	1.4	196714	3	US-09-949-016-15474	Sequence 15474, A
C 427	43	1.4	237510	3	US-09-949-016-14273	Sequence 14273, A
C 428	43	1.4	346112	3	US-09-949-016-13165	Sequence 13165, A
C 429	43	1.4	336	3	US-09-385-982-17	Sequence 17, App1
C 430	43	1.4	601	3	US-09-949-016-19618	Sequence 19618, A
C 431	43	1.4	601	3	US-09-949-016-25518	Sequence 25518, A
C 432	43	1.4	601	3	US-09-949-016-28910	Sequence 28910, A
C 433	43	1.4	601	3	US-09-949-016-31419	Sequence 31419, A
C 434	43	1.4	601	3	US-09-949-016-37684	Sequence 37684, A
C 435	43	1.4	601	3	US-09-949-016-37704	Sequence 37704, A
C 436	43	1.4	601	3	US-09-949-016-3927	Sequence 3927, A
C 437	43	1.4	601	3	US-09-949-016-45798	Sequence 45798, A
C 438	43	1.4	601	3	US-09-949-016-45420	Sequence 45420, A
C 439	43	1.4	601	3	US-09-949-016-57030	Sequence 57030, A
C 440	43	1.4	601	3	US-09-949-016-57821	Sequence 57821, A
C 441	43	1.4	601	3	US-09-949-016-57822	Sequence 57822, A
C 442	43	1.4	601	3	US-09-949-016-59688	Sequence 59688, A
C 443	43	1.4	601	3	US-09-949-016-73696	Sequence 73696, A
C 444	43	1.4	601	3	US-09-949-016-86370	Sequence 86370, A
C 445	43	1.4	601	3	US-09-949-016-86370	Sequence 86370, A
C 446	43	1.4	601	3	US-09-949-016-105326	Sequence 105326, A
C 447	43	1.4	601	3	US-09-949-016-121736	Sequence 121736, A
C 448	43	1.4	601	3	US-09-949-016-12587	Sequence 12587, A
C 449	43	1.4	601	3	US-09-949-016-125964	Sequence 125964, A
C 450	43	1.4	601	3	US-09-949-016-127231	Sequence 127231, A
C 451	43	1.4	601	3	US-09-949-016-127826	Sequence 127826, A
C 452	43	1.4	601	3	US-09-949-016-128163	Sequence 128163, A
C 453	43	1.4	601	3	US-09-949-016-135139	Sequence 135139, A
C 454	43	1.4	601	3	US-09-949-016-139090	Sequence 139090, A
C 455	43	1.4	601	3	US-09-949-016-141088	Sequence 141088, A
C 456	43	1.4	601	3	US-09-949-016-146571	Sequence 146571, A
C 457	43	1.4	601	3	US-09-949-016-147357	Sequence 147357, A
C 458	43	1.4	601	3	US-09-949-016-169230	Sequence 169230, A
C 459	43	1.4	601	3	US-09-949-016-171811	Sequence 171811, A
C 460	43	1.4	601	3	US-09-949-016-171812	Sequence 171812, A
C 461	43	1.4	601	3	US-09-949-016-175700	Sequence 175700, A
C 462	43	1.4	601	3	US-09-949-016-176743	Sequence 176743, A

C 463	42	1.3	601	3	US-09-949-016-177625	Sequence 177625, A
C 464	42	1.3	601	3	US-09-949-016-180839	Sequence 180839, A
C 465	42	1.3	601	3	US-09-949-016-186854	Sequence 186854, A
C 466	42	1.3	601	3	US-09-949-016-188045	Sequence 188045, A
C 467	42	1.3	601	3	US-09-949-016-188074	Sequence 188074, A
C 468	42	1.3	601	3	US-09-949-016-188551	Sequence 188551, A
C 469	42	1.3	601	3	US-09-949-016-199254	Sequence 199254, A
C 470	42	1.3	601	3	US-09-949-016-199255	Sequence 199255, A
C 471	42	1.3	601	3	US-09-949-002-1276	Sequence 1276, App
C 472	42	1.3	601	3	US-09-949-002-1277	Sequence 1277, App
C 473	42	1.3	601	3	US-09-949-002-10408	Sequence 10408, A
C 474	42	1.3	601	3	US-09-949-002-10409	Sequence 10409, A
C 475	42	1.3	5227	3	US-09-973-278-919	Sequence 919, App
C 476	42	1.3	5235	3	US-09-973-278-918	Sequence 918, App
C 477	42	1.3	5969	3	US-09-949-016-1342	Sequence 1342, App
C 478	42	1.3	5970	3	US-09-949-016-938	Sequence 938, App
C 479	42	1.3	6328	3	US-09-949-016-14847	Sequence 14847, A
C 480	42	1.3	6831	3	US-09-949-016-12281	Sequence 12281, A
C 481	42	1.3	7584	3	US-09-949-016-11836	Sequence 11836, A
C 482	42	1.3	7585	3	US-09-949-016-16737	Sequence 16737, A
C 483	42	1.3	9900	3	US-09-949-016-13392	Sequence 13392, A
C 484	42	1.3	11022	3	US-09-949-016-15502	Sequence 15502, A
C 485	42	1.3	15661	3	US-09-949-016-14031	Sequence 14031, A
C 486	42	1.3	15664	3	US-09-949-016-13030	Sequence 13030, A
C 487	42	1.3	18264	3	US-09-949-016-14465	Sequence 14465, A
C 488	42	1.3	18605	3	US-09-949-016-12223	Sequence 12223, A
C 489	42	1.3	20059	3	US-09-949-016-14716	Sequence 14716, A
C 490	42	1.3	20060	3	US-09-949-016-13055	Sequence 13055, A
C 491	42	1.3	23995	3	US-09-949-016-17363	Sequence 17363, A
C 492	42	1.3	23998	3	US-09-949-016-14009	Sequence 14009, A
C 493	42	1.3	24841	3	US-09-949-016-16056	Sequence 16056, A
C 494	42	1.3	24847	3	US-09-949-016-14682	Sequence 14682, A
C 495	42	1.3	27592	3	US-09-949-016-13055	Sequence 13055, A
C 496	42	1.3	27667	3	US-09-949-016-13371	Sequence 13371, A
C 497	42	1.3	27923	3	US-09-949-016-15396	Sequence 15396, A
C 498	42	1.3	28791	3	US-09-949-016-12437	Sequence 12437, A
C 499	42	1.3	28823	3	US-09-949-016-15879	Sequence 15879, A
C 500	42	1.3	28823	3	US-09-949-016-15879	Sequence 15879, A

## ALIGNMENTS

RESULT 1  
US-09-513-999C-3793/C  
; Sequence 3793, Application US/09513999C  
; Patent No. 6783961  
; GENERAL INFORMATION:  
; APPLICANT: Dumas Milne Edwards, J.B.  
; APPLICANT: Duclet, A.  
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.  
; Patent No. 6783961  
; FILE REFERENCE: 59, US2, REG  
; CURRENT APPLICATION NUMBER: US/09/513, 999C  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/122, 487  
; NUMBER OF SEQ ID NOS: 36681  
; SEQ ID NO 3793  
; LENGTH: 381  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURES:  
; NAME/KEY: CDS  
; LOCATION: 12..347  
; FEATURE:  
; NAME/KEY: sig\_peptide  
; LOCATION: 12..158  
; OTHER INFORMATION: score 7.7  
; OTHER INFORMATION: seq LC1LSSDSPA/50  
; FEATURE:



```

; NAME/KEY: misc_feature
; LOCATION: 322
; OTHER INFORMATION: s=g or c
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 370
; OTHER INFORMATION: k=g or t
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 376
; OTHER INFORMATION: r=a or g
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: 55
; OTHER INFORMATION: Xaa=Ser or Thr
US-09-513-999C-3793

Query Match          1.7%; Score 52; DB 3; Length 381;
Best Local Similarity 100.0%; Pred. No. 3.5e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTCACCTGCACTCCAGCTGGGCAACAGACAAGACTCTGTCTC 3122
DB 107 CAAGATTGTCACCTGCACTCCAGCTGGGCAACAGACAAGACTCTGTCTC 56

RESULT 2
US-09-949-016-72610/C
; Sequence 72610, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72610
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-72610

Query Match          1.7%; Score 52; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 3.4e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTCACCTGCACTCCAGCTGGGCAACAGACAAGACTCTGTCTC 3122
DB 71 CAAGATTGTCACCTGCACTCCAGCTGGGCAACAGACAAGACTCTGTCTC 20

RESULT 3
US-09-949-016-75710
; Sequence 75710, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75710
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-75710

Query Match          1.7%; Score 52; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 3.4e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTCACCTGCACTCCAGCTGGGCAACAGACAAGACTCTGTCTC 3122
DB 71 CAAGATTGTCACCTGCACTCCAGCTGGGCAACAGACAAGACTCTGTCTC 20

RESULT 4
US-09-949-016-75723
; Sequence 75723, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75723
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-75723

Query Match          1.7%; Score 52; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 3.4e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCAAGTGTGATCACTGAGGCCAGGAGTTGAGACCAAGCCTGGCCAAACAT 2940
DB 191 GAGGCAAGTGTGATCACTGAGGCCAGGAGTTGAGACCAAGCCTGGCCAAACAT 242

RESULT 5
US-09-949-016-13931/C
; Sequence 13931, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13931
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13931/C

Query Match          1.7%; Score 52; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 3.4e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCAAGTGTGATCACTGAGGCCAGGAGTTGAGACCAAGCCTGGCCAAACAT 2940
DB 191 GAGGCAAGTGTGATCACTGAGGCCAGGAGTTGAGACCAAGCCTGGCCAAACAT 242
```

```

; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75710
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-75710

Query Match          1.7%; Score 52; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 3.4e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCAAGTGTGATCACTGAGGCCAGGAGTTGAGACCAAGCCTGGCCAAACAT 2940
DB 191 GAGGCAAGTGTGATCACTGAGGCCAGGAGTTGAGACCAAGCCTGGCCAAACAT 242

RESULT 4
US-09-949-016-75723
; Sequence 75723, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 75723
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-75723

Query Match          1.7%; Score 52; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 3.4e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCAAGTGTGATCACTGAGGCCAGGAGTTGAGACCAAGCCTGGCCAAACAT 2940
DB 191 GAGGCAAGTGTGATCACTGAGGCCAGGAGTTGAGACCAAGCCTGGCCAAACAT 242

RESULT 5
US-09-949-016-13931/C
; Sequence 13931, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13931
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-13931/C

Query Match          1.7%; Score 52; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 3.4e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCAAGTGTGATCACTGAGGCCAGGAGTTGAGACCAAGCCTGGCCAAACAT 2940
DB 191 GAGGCAAGTGTGATCACTGAGGCCAGGAGTTGAGACCAAGCCTGGCCAAACAT 242
```

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 13931  
LENGTH: 15415  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-13931

Query Match 1.7%; Score 52; DB 3; Length 15415;  
Best Local Similarity 100.0%; Pred.No. 3.1e-14;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGAGGTGATCCTGAGGCGCAGAGTTCAGACGAGCCTGGCCAAAT 2940  
DB 11938 GAGGAGGTGATCCTGAGGCGCAGAGTTCAGACGAGCCTGGCCAAAT 11887

RESULT 6  
US-09-949-016-13932/c

Sequence 13932, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: C1001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 13932

LENGTH: 15415

TYPE: DNA

ORGANISM: Human

US-09-949-016-13932

Query Match 1.7%; Score 52; DB 3; Length 15415;  
Best Local Similarity 100.0%; Pred.No. 3.1e-14;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGAGGTGATCCTGAGGCGCAGAGTTCAGACGAGCCTGGCCAAAT 2940  
DB 11938 GAGGAGGTGATCCTGAGGCGCAGAGTTCAGACGAGCCTGGCCAAAT 11887

RESULT 7  
US-09-949-016-1397/c

Sequence 1397, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: C1001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 1397

LENGTH: 25969

TYPE: DNA

ORGANISM: Human

US-09-949-016-1397

US-09-949-016-13397

Query Match 1.7%; Score 52; DB 3; Length 25969;  
Best Local Similarity 100.0%; Pred.No. 3e-14;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGTTGTGCACTGCACTCAGGCGCAAGAGCAAGACTGTCTC 3122  
DB 22148 CAAGTTGTGCACTGCACTCAGGCGCAAGAGCAAGACTGTCTC 22097

RESULT 8  
US-09-949-016-13220/c

Sequence 13220, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: C1001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 13220

LENGTH: 77586

TYPE: DNA

ORGANISM: Human

US-09-949-016-13220

Query Match 1.7%; Score 52; DB 3; Length 77586;  
Best Local Similarity 100.0%; Pred.No. 2.9e-14;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGTTGTGCACTGCACTCAGGCGCAAGAGCAAGACTGTCTC 3122  
DB 43078 CAAGTTGTGCACTGCACTCAGGCGCAAGAGCAAGACTGTCTC 43027

RESULT 9  
US-09-949-016-13221/c

Sequence 13221, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: C1001307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 13221

LENGTH: 77586

TYPE: DNA

ORGANISM: Human

US-09-949-016-13221

Query Match 1.7%; Score 52; DB 3; Length 77586;  
Best Local Similarity 100.0%; Pred.No. 2.9e-14;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGCGCACTGCACTCCAGCGCTGGGCAACAGAGCAAGACTGTCTC 3122  
DB 43078 CAGGATTGGCCACTGCACTCCAGCGCTGGGCAACAGAGCAAGACTGTCTC 43027

## RESULT 10

US-09-918-686-1  
; Sequence 1, Application US/09918686  
; Patent No. 6475739  
; GENERAL INFORMATION:  
; APPLICANT: Brunkow, Mary  
; APPLICANT: Prohl, Sean  
; APPLICANT: Paepert, Bryan  
; APPLICANT: Staehling-Hampton, Karen  
; TITLE OF INVENTION: METHODS FOR IDENTIFYING  
; FILE REFERENCE: 240083.515  
; CURRENT APPLICATION NUMBER: US/09/918,686  
; PRIOR FILING DATE: 2001-07-30  
; NUMBER OF SEQ ID NOS: 105  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 92139  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: 7043, 8369, 8401  
; OTHER INFORMATION: n = A,T,C or G  
US-09-918-686-1

Query Match 1.7%; Score 52; DB 3; Length 92139;  
Best Local Similarity 100.0%; Pred. No. 2.9e-14;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGAGGTGATCACTGAGGCGCAGAGTTGAGACCGAGCTGGCCAAAT 2940  
DB 85294 GAGGAGGTGATCACTGAGGCGCAGAGTTGAGACCGAGCTGGCCAAAT 85345

## RESULT 11

US-09-949-016-15787  
; Sequence 15787, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15787  
; LENGTH: 120727  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)... (120727)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-15787

Query Match 1.7%; Score 52; DB 3; Length 120727;  
Best Local Similarity 100.0%; Pred. No. 2.9e-14;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGAGGTGATCACTGAGGCGCAGAGTTGAGACCGAGCTGGCCAAAT 2940  
DB 58522 GAGGAGGTGATCACTGAGGCGCAGAGTTGAGACCGAGCTGGCCAAAT 58573

## RESULT 12

US-09-949-016-15788  
; Sequence 15788, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15788  
; LENGTH: 120727  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)... (120727)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-15788

Query Match 1.7%; Score 52; DB 3; Length 120727;  
Best Local Similarity 100.0%; Pred. No. 2.9e-14;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGAGGTGATCACTGAGGCGCAGAGTTGAGACCGAGCTGGCCAAAT 2940  
DB 58522 GAGGAGGTGATCACTGAGGCGCAGAGTTGAGACCGAGCTGGCCAAAT 58573

## RESULT 13

US-09-949-016-13870  
; Sequence 13870, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13870  
; LENGTH: 155266  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-13870

Query Match 1.7%; Score 52; DB 3; Length 155266;  
Best Local Similarity 100.0%; Pred. No. 2.9e-14;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy	3071	CAAGATTGTGGCACCTGCACCTCAGCCCTGGGCAACAGACAAGACTTGTCTC	3122
D6	139149	CAAGATTGTGGCACCTGCACCTCAGCCCTGGGCAACAGACAAGACTTGTCTC	139200

```

RESULT 14
US-09-949-016-16873/c
: Sequence 16873, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CLO01107
: CURRENT APPLICATION NUMBER: US/09/949,016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 16873
: LENGTH: 35688
: TYPE: DNA
: ORGANISM: Human
US-09-949-016-16873

```

```

RESULT 15
US-09-949-016-12227
; Sequence 12227, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL0010307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12227
; LENGTH: 156942
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(156942)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12227

```

Db 118483 AGCAGTTGGATCACCTGAGGCCAGAGTTCGAGACCAAGCTTGCCACAT 118533

```

RESULT 16
US-09-949-016-15946
? Sequence 15946, Application US/09949016
? Patent No. 681233
? GENERAL INFORMATION:
? APPLICANT: VENTER, J. Craig et al.
? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: CL001307
? CURRENT APPLICATION NUMBER: US/09/949,016
? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 15946
? LENGTH: 156950
? TYPE: DNA
? ORGANISM: Human
? FEATURE:
? NAME/KEY: misc_feature
? LOCATION: (1)..(156950)
? OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15946

```

```

RESULT 17
US-09-513-999C-28539/C
Sequence 28539, Application US/09513999C
Patent No. 6783961
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Duclert, A.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
Patent No. 6783961
FILE REFERENCE: 59, US2, REG
CURRENT APPLICATION NUMBER: US/09/513, 999C
CURRENT FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/122,487
PRIOR FILING DATE: 1999-02-26
NUMBER OF SEQ ID NOS: 36681
SOFTWARE: Patent.pm
SEQ ID NO 28539
LENGTH: 255
TYPE: DNA
ORGANISM: Homo sapiens
US-09-513-999C-28539

```

RESULT 18  
US-09-973-2278-870  
Sequence 870, Application US/09973278  
Patent No. 6924354  
GENERAL INFORMATION:  
APPLICANT: Fischer et al.  
TITLE OF INVENTION: 123 Human Secreted Proteins  
FILE REFERENCE: P2010P2  
CURRENT APPLICATION NUMBER: US/09/973,278  
PRIORITY FILING DATE: 2001-10-10  
PRIOR APPLICATION NUMBER: 60/239,899  
PRIOR FILING DATE: 2000-10-13  
PRIOR APPLICATION NUMBER: 09/227,357  
PRIOR FILING DATE: 1999-01-08  
PRIOR APPLICATION NUMBER: PCT/US98/13684  
PRIOR FILING DATE: 1998-07-07  
PRIOR APPLICATION NUMBER: 60/051,926  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/052,793  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/051,925  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/051,929  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/052,803  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/052,732  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/051,931  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/051,932  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/051,916  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/051,930  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/051,918  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/051,920  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/052,733  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/052,795  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/051,919  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/051,928  
PRIOR FILING DATE: 1997-07-08  
PRIOR APPLICATION NUMBER: 60/055,722  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/055,723  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/055,948  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/055,949  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/055,953  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/055,950  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/055,947  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/055,964  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/056,360  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/055,684  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/055,984  
PRIOR FILING DATE: 1997-08-18  
PRIOR APPLICATION NUMBER: 60/055,954  
PRIOR FILING DATE: 1997-08-18

PRIOR APPLICATION NUMBER: 60/058,785  
PRIOR FILING DATE: 1997-09-12  
PRIOR APPLICATION NUMBER: 60/058,664  
PRIOR FILING DATE: 1997-09-12  
PRIOR APPLICATION NUMBER: 60/058,660  
PRIOR FILING DATE: 1997-09-12  
PRIOR APPLICATION NUMBER: 60/058,661  
PRIOR FILING DATE: 1997-09-12  
NUMBER OF SEQ ID NOS: 947  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 870  
LENGTH: 288  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-973-278-870  
Query Match  
Best Local Similarity 1.6%; Score 50; DB 3; Length 288;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 3073 AGATTGTGCGACTGCACTCCAGCTGGGCAAGAGCAAGACTCTGTCTC 3122  
DB 221 AGATTGTGCGACTGCACTCCAGCTGGGCAAGAGCAAGACTCTGTCTC 270  
RESULT 19  
US-09-949-016-47563  
Sequence 47563, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
PRIORITY FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 47563  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-47563  
Query Match  
Best Local Similarity 1.6%; Score 50; DB 3; Length 601;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
CY 3073 AGATTGTGCGACTGCACTCCAGCTGGGCAAGAGCAAGACTCTGTCTC 3122  
DB 531 AGATTGTGCGACTGCACTCCAGCTGGGCAAGAGCAAGACTCTGTCTC 580  
RESULT 20  
US-09-949-016-125579  
Sequence 125579, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
PRIORITY FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768

```

; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 125579
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-125579

Query Match
Best Local Similarity 1.6%; Score 50; DB 3; Length 601;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACAGACTCTGTCTC 3122
DB 221 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACAGACTCTGTCTC 270

RESULT 21
US-09-949-016-125580
; Sequence 125580, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 125580
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-125580

Query Match
Best Local Similarity 1.6%; Score 50; DB 3; Length 601;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACAGACTCTGTCTC 3122
DB 203 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACAGACTCTGTCTC 252

RESULT 22
US-09-949-016-125581
; Sequence 125581, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; TYPE: DNA
```

```

; SEQ ID NO: 125581
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-125581

Query Match
Best Local Similarity 1.6%; Score 50; DB 3; Length 601;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACAGACTCTGTCTC 3122
DB 46 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACAGACTCTGTCTC 95

RESULT 23
US-09-949-016-163881
; Sequence 163881, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 163881
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-163881

Query Match
Best Local Similarity 1.6%; Score 50; DB 3; Length 601;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACAGACTCTGTCTC 3122
DB 179 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACAGACTCTGTCTC 228

RESULT 24
US-09-544-398B-10
; Sequence 10, Application US/09544398B
; Patent No. 6770461
; GENERAL INFORMATION:
; APPLICANT: Carulli, John P.
; APPLICANT: Little, Randall D.
; APPLICANT: Recker, Robert R.
; APPLICANT: Johnson, Mark L.
; TITLE OF INVENTION: High bone mass gene of 11q13.3
; FILE REFERENCE: 032796-013
; CURRENT APPLICATION NUMBER: US/09/544,398B
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 641
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 10
; LENGTH: 8705
; TYPE: DNA
```

ORGANISM: Homo sapiens  
US-09-544-398B-10

Query Match 1.6%; Score 50; DB 3; Length 8705;  
Best Local Similarity 100.0%; Pred. No. 2.7e-13;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCGGGCAACAGAGCAAGACTGTCTC 3122  
DB 6492 AGATTGTGCACCTGCACCTCGGGCAACAGAGCAAGACTGTCTC 6541

## RESULT 25

US-09-543-771B-10  
Sequence 10, Application US/09543771B  
Patent No. 6780609  
GENERAL INFORMATION:  
APPLICANT: Carulli, John P.  
APPLICANT: Little, Randall D.  
APPLICANT: Becker, Robert R.  
APPLICANT: Johnson, Mark L.  
TITLE OF INVENTION: High bone mass gene of 11q13.3  
FILE REFERENCE: 032796-014  
CURRENT APPLICATION NUMBER: US/09/543,771B  
CURRENT FILING DATE: 2000-04-05  
PRIOR APPLICATION NUMBER: US 09/229,319  
PRIOR FILING DATE: 1999-01-13  
PRIOR APPLICATION NUMBER: US 60/071,449  
PRIOR FILING DATE: 1998-01-13  
PRIOR APPLICATION NUMBER: US 60/105,511  
PRIOR FILING DATE: 1998-10-23  
NUMBER OF SEQ ID NOS: 641  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 10  
LENGTH: 8705  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-543-771B-10

Query Match 1.6%; Score 50; DB 3; Length 8705;  
Best Local Similarity 100.0%; Pred. No. 2.7e-13;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCGGGCAACAGAGCAAGACTGTCTC 3122  
DB 6492 AGATTGTGCACCTGCACCTCGGGCAACAGAGCAAGACTGTCTC 6541

## RESULT 26

US-09-949-016-15304  
Sequence 15304, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15304  
LENGTH: 36954  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-15304

Query Match 1.6%; Score 50; DB 3; Length 36954;  
Best Local Similarity 100.0%; Pred. No. 2.6e-13;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCGGGCAACAGAGCAAGACTGTCTC 3122  
DB 27720 AGATTGTGCACCTGCACCTCGGGCAACAGAGCAAGACTGTCTC 27769

## RESULT 27

US-09-949-016-12384  
Sequence 12384, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12384  
LENGTH: 39154  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(39154)  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12384

Query Match 1.6%; Score 50; DB 3; Length 39154;  
Best Local Similarity 100.0%; Pred. No. 2.6e-13;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCGGGCAACAGAGCAAGACTGTCTC 3122  
DB 38617 AGATTGTGCACCTGCACCTCGGGCAACAGAGCAAGACTGTCTC 38666

## RESULT 28

US-09-949-016-12801  
Sequence 12801, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12801  
LENGTH: 39154  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(39154)



OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12801

Query Match 1.6%; Score 50; DB 3; Length 39154;  
Best Local Similarity 100.0%; Pred. No. 2.6e-13;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
DB 38617 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 38666

RESULT 29  
US-09-949-016-14326  
Sequence 14326, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14326  
LENGTH: 39443  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1) - (39443)  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14326

Query Match 1.6%; Score 50; DB 3; Length 39443;  
Best Local Similarity 100.0%; Pred. No. 2.6e-13;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
DB 38617 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 38666

RESULT 30  
US-09-949-016-14327  
Sequence 14327, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14327  
LENGTH: 39443  
TYPE: DNA  
ORGANISM: Human

FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1) - (39443)  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14327

Query Match 1.6%; Score 50; DB 3; Length 39443;  
Best Local Similarity 100.0%; Pred. No. 2.6e-13;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
DB 38617 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 38666

RESULT 31  
US-09-949-016-17153  
Sequence 17153, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 17153  
LENGTH: 76401  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-17153

Query Match 1.6%; Score 50; DB 3; Length 76401;  
Best Local Similarity 100.0%; Pred. No. 2.5e-13;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
DB 56563 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 56612

RESULT 32  
US-09-949-016-16335  
Sequence 16335, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 16335  
LENGTH: 87190  
TYPE: DNA  
ORGANISM: Human  
FEATURE:

NAME/KEY: misc\_feature  
LOCATION: (1)...(87190)  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-16335

Query Match 1.6%; Score 50; DB 3; Length 87190;  
Best Local Similarity 100.0%; Pred. No. 2.5e-13;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122  
DB 57590 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 57639

RESULT 33  
US-09-949-016-13139/C  
Sequence 13139, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 13139  
LENGTH: 95648  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(95648)  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-13139

Query Match 1.6%; Score 50; DB 3; Length 95648;  
Best Local Similarity 100.0%; Pred. No. 2.5e-13;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122  
DB 68111 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 68062

RESULT 34  
US-09-949-016-14199  
Sequence 14199, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14199  
LENGTH: 140844

TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(140844)  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14199

Query Match 1.6%; Score 50; DB 3; Length 140844;  
Best Local Similarity 100.0%; Pred. No. 2.5e-13;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122  
DB 140232 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 140281

RESULT 35  
US-09-949-016-12881  
Sequence 12881, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12881  
LENGTH: 197336  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(197336)  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12881

Query Match 1.6%; Score 50; DB 3; Length 197336;  
Best Local Similarity 100.0%; Pred. No. 2.5e-13;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122  
DB 189434 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 189483

RESULT 36  
US-09-949-016-14376  
Sequence 14376, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14376
; LENGTH: 197337
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)_(197337)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-14376

Query Match
Best Local Similarity 100.0%; Score 50; DB 3; Length 197337;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122
DB 189434 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 189483

RESULT 37
US-09-949-002-738
; Sequence 738, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 738
; LENGTH: 197337
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(197337)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-002-738

Query Match
Best Local Similarity 100.0%; Score 50; DB 3; Length 197337;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122
DB 189434 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 189483

RESULT 38
US-09-949-002-687
; Sequence 687, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 687
; LENGTH: 234287

; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(234287)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-002-687

Query Match
Best Local Similarity 100.0%; Score 50; DB 3; Length 234287;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122
DB 189434 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 189483

RESULT 39
US-09-949-016-17272
; Sequence 17272, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17272
; LENGTH: 234288
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(234288)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17272

Query Match
Best Local Similarity 100.0%; Score 50; DB 3; Length 234288;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122
DB 189434 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 189483

RESULT 40
US-09-949-002-841
; Sequence 841, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 841
; LENGTH: 234288
```

```

; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(234288)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-002-841
```

```

Query Match          1.6%; Score 50; DB 3; Length 234288;
Best Local Similarity 100.0%; Pred.No. 2.4e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3073 AGATTGTCACCTGACCTCAGGCTGGGCAACAGACCAAGCTGTCTC 3122
          |||||
DB      189434 AATATTGTCACCTGACCTCAGGCTGGGCAACAGACCAAGCTGTCTC 189483
```

```

RESULT 41
US-09-949-016-12583/C
; Sequence 12583, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12583
; LENGTH: 254964
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(254964)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12583
```

```

Query Match          1.6%; Score 50; DB 3; Length 254964;
Best Local Similarity 100.0%; Pred.No. 2.4e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2888 TGAGGCGAGTGATCACTCAGGCTGGGCAACAGACCAAGCTGTCTC 2937
          |||||
DB      108004 TGAGGCGAGTGATCACTCAGGCTGGGCAACAGACCAAGCTGTCTC 107955
```

```

RESULT 42
US-09-949-016-17392/C
; Sequence 17392, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17392
; LENGTH: 254964
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(254964)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-17392
```

```

Query Match          1.6%; Score 50; DB 3; Length 254964;
Best Local Similarity 100.0%; Pred.No. 2.4e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2888 TGAGGCGAGTGATCACTCAGGCTGGGCAACAGACCAAGCTGTCTC 2937
          |||||
DB      108004 TGAGGCGAGTGATCACTCAGGCTGGGCAACAGACCAAGCTGTCTC 107955
```

```

RESULT 43
US-09-949-016-164699/C
; Sequence 164699, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 164699
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(601)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-164699
```

```

Query Match          1.5%; Score 47; DB 3; Length 601;
Best Local Similarity 100.0%; Pred.No. 7.5e-12;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3073 AGATTGTCACCTGACCTCAGGCTGGGCAACAGACCAAGCTGTCTC 3119
          |||||
DB      580 AATATTGTCACCTGACCTCAGGCTGGGCAACAGACCAAGCTGTCTC 534
```

```

RESULT 44
US-09-949-016-16733
; Sequence 16733, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
```

SEQ ID NO 16733  
LENGTH: 12880  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-16733

Query Match 1.5%; Score 47; DB 3; Length 12880;  
Best Local Similarity 100.0%; Pred. No. 6.5e-12;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2888 TGAGGAGGTGATGATCCTGAGGCGCAGAGTTGAGAGCCAGCTGGC 2934  
DB 5157 TGAGGAGGTGATGATCCTGAGGCGCAGAGTTGAGAGCCAGCTGGC 5203

RESULT 45  
US-09-949-016-14330/c  
Sequence 14330, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
PRIOR FILING DATE: 2000-04-14  
PRIOR FILING DATE: 2000-10-20  
PRIOR FILING DATE: 2000-10-20  
PRIOR FILING DATE: 2000-10-03  
PRIOR FILING DATE: 2000-10-03  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14330  
LENGTH: 45833  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-14330

Query Match 1.5%; Score 47; DB 3; Length 45833;  
Best Local Similarity 100.0%; Pred. No. 6.5e-12;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAGATTGTGCGCACTGCACTCCAGCTGGGCAACAGCAAGACTCT 3117  
DB 22968 CAGATTGTGCGCACTGCACTCCAGCTGGGCAACAGCAAGACTCT 22922

RESULT 46  
US-09-949-016-14331/c  
Sequence 14331, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
PRIOR FILING DATE: 2000-04-14  
PRIOR FILING DATE: 2000-10-20  
PRIOR FILING DATE: 2000-10-20  
PRIOR FILING DATE: 2000-10-03  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14331  
LENGTH: 45833  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-14331

Query Match 1.5%; Score 47; DB 3; Length 45833;  
Best Local Similarity 100.0%; Pred. No. 6.5e-12;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAGATTGTGCGCACTGCACTCCAGCTGGGCAACAGCAAGACTCT 3117  
DB 22968 CAGATTGTGCGCACTGCACTCCAGCTGGGCAACAGCAAGACTCT 22922

RESULT 47  
US-09-949-016-14332/c  
Sequence 14332, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
PRIOR FILING DATE: 2000-04-14  
PRIOR FILING DATE: 2000-10-20  
PRIOR FILING DATE: 2000-10-20  
PRIOR FILING DATE: 2000-10-03  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14332  
LENGTH: 45833  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-14332

Query Match 1.5%; Score 47; DB 3; Length 45833;  
Best Local Similarity 100.0%; Pred. No. 6.5e-12;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAGATTGTGCGCACTGCACTCCAGCTGGGCAACAGCAAGACTCT 3117  
DB 22968 CAGATTGTGCGCACTGCACTCCAGCTGGGCAACAGCAAGACTCT 22922

RESULT 48  
US-09-949-016-14333/c  
Sequence 14333, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
PRIOR FILING DATE: 2000-04-14  
PRIOR FILING DATE: 2000-10-20  
PRIOR FILING DATE: 2000-10-20  
PRIOR FILING DATE: 2000-10-03  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14333  
LENGTH: 45833  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-14333

Query Match 1.5%; Score 47; DB 3; Length 45833;  
Best Local Similarity 100.0%; Pred. No. 6.5e-12;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGGCACTGCATCCAGCTGGGCAACAGAGAAGACTCT 3117  
|||||  
Db 22968 CAAGATTGGCACTGCATCCAGCTGGGCAACAGAGAAGACTCT 22922

RESULT 49  
US-09-949-016-16519

; Sequence 16519, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 16519

; LENGTH: 67386

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(67386)

; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-16519

Query Match

Best Local Similarity 1.5%; Score 47; DB 3; Length 67386;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3076 TTGTGCACTGCATCTCCAGCTGGGCAACAGAGAAGACTCTCTC 3122

|||||

Db 49989 TTGTGCACTGCATCTCCAGCTGGGCAACAGAGAAGACTCTCTC 50035

|||||

RESULT 50

US-09-949-016-16366/C

; Sequence 16366, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 16366

; LENGTH: 108310

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(108310)

; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-16366

Query Match

Best Local Similarity 100.0%; Pred. No. 6.4e-12;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGGCACTGCATCTCCAGCTGGGCAACAGAGAAGACTCTGT 3119

|||||

Db 63285 AGATTGGCACTGCATCTCCAGCTGGGCAACAGAGAAGACTCTGT 63239

|||||

RESULT 51

US-09-949-016-17030

; Sequence 17030, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 17030

; LENGTH: 139049

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(139049)

; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-17030

Query Match

Best Local Similarity 1.5%; Score 47; DB 3; Length 139049;

Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGGCACTGCATCTCCAGCTGGGCAACAGAGAAGACTCTGT 3119

|||||

Db 42519 AGATTGGCACTGCATCTCCAGCTGGGCAACAGAGAAGACTCTGT 42565

|||||

RESULT 52

US-09-949-016-65112

; Sequence 65112, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; CURRENT FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 65112

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

; FEATURE:

; NAME/KEY: misc\_feature

; LOCATION: (1)...(601)

; OTHER INFORMATION: n = A,T,C or G

US-09-949-016-65112

Query Match

Best Local Similarity 1.5%; Score 46; DB 3; Length 601;

Matches 47; Conservative 100.0%; Pred. No. 2.2e-11;

Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAACAT 2940

Db 67 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAACAT 112

RESULT 53

US-09-949-016-65113

; Sequence 65113, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 65113

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-65113

Query Match

Best Local Similarity 1.5%; Score 46; DB 3; Length 601;

Best Local Similarity 100.0%; Pred. No. 2.2e-11;

Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAACAT 2940

Db 395 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAACAT 440

RESULT 54

US-09-949-016-65114

; Sequence 65114, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 65114

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-65114

Query Match

Best Local Similarity 1.5%; Score 46; DB 3; Length 601;

Best Local Similarity 100.0%; Pred. No. 2.2e-11;

Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAACAT 2940

Db 406 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAACAT 451

RESULT 55

US-09-949-016-137514

; Sequence 137514, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 137514

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-137514

Query Match

Best Local Similarity 1.5%; Score 46; DB 3; Length 601;

Best Local Similarity 100.0%; Pred. No. 2.2e-11;

Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGGCCACTGCACTCCAGCTGGGCAAGAGCAAGATC 3116

Db 334 CAAGATTGGCCACTGCACTCCAGCTGGGCAAGAGCAAGATC 369

RESULT 56

US-09-949-016-155725/c

; Sequence 155725, Application US/09949016

; Patent No. 6812339

; GENERAL INFORMATION:

; APPLICANT: VENTER, J. Craig et al.

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

; FILE REFERENCE: CL001307

; CURRENT APPLICATION NUMBER: US/09/949,016

; PRIOR FILING DATE: 2000-04-14

; PRIOR APPLICATION NUMBER: 60/241,755

; PRIOR FILING DATE: 2000-10-20

; PRIOR APPLICATION NUMBER: 60/237,768

; PRIOR FILING DATE: 2000-10-03

; PRIOR APPLICATION NUMBER: 60/231,498

; PRIOR FILING DATE: 2000-09-08

; NUMBER OF SEQ ID NOS: 207012

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 155725

; LENGTH: 601

; TYPE: DNA

; ORGANISM: Human

US-09-949-016-155725

Query Match

Best Local Similarity 1.5%; Score 46; DB 3; Length 601;

Best Local Similarity 100.0%; Pred. No. 2.2e-11;

Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAACAT 2940

Db 486 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAACAT 441

RESULT 57

US-09-949-016-155726/c

; Sequence 155726, Application US/09949016



```

; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 155726
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-155726

Query Match          1.5%; Score 46; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.2e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2895 GGTGATTCACCTGAGCCGAGATTGGAACCCAGCTGGCCCAACAT 2340
DB      477 GGTGATTCACCTGAGCCGAGATTGGAACCCAGCTGGCCCAACAT 432

RESULT 58
US-09-949-016-169304/c
; Sequence 169304, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 169304
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-169304

Query Match          1.5%; Score 46; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.2e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3071 CAAGATTGTGCGACTGCACTCCAGCTGGGCAACAGCAAGACTC 3116
DB      468 CAAGATTGTGCGACTGCACTCCAGCTGGGCAACAGCAAGACTC 423

RESULT 59
US-09-949-016-169305/c
; Sequence 169305, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 169305
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-169305

Query Match          1.5%; Score 46; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.2e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3071 CAAGATTGTGCGACTGCACTCCAGCTGGGCAACAGCAAGACTC 3116
DB      211 CAAGATTGTGCGACTGCACTCCAGCTGGGCAACAGCAAGACTC 166

RESULT 60
US-09-949-016-169306/c
; Sequence 169306, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 169306
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-169306

Query Match          1.5%; Score 46; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.2e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3071 CAAGATTGTGCGACTGCACTCCAGCTGGGCAACAGCAAGACTC 3116
DB      211 CAAGATTGTGCGACTGCACTCCAGCTGGGCAACAGCAAGACTC 166

RESULT 61
US-09-949-016-198268
; Sequence 198268, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
```

```

; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 169305
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-169305

Query Match          1.5%; Score 46; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.2e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3071 CAAGATTGTGCGACTGCACTCCAGCTGGGCAACAGCAAGACTC 3116
DB      353 CAAGATTGTGCGACTGCACTCCAGCTGGGCAACAGCAAGACTC 308

RESULT 60
US-09-949-016-169306/c
; Sequence 169306, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 169306
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-169306

Query Match          1.5%; Score 46; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 2.2e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3071 CAAGATTGTGCGACTGCACTCCAGCTGGGCAACAGCAAGACTC 3116
DB      211 CAAGATTGTGCGACTGCACTCCAGCTGGGCAACAGCAAGACTC 166

RESULT 61
US-09-949-016-198268
; Sequence 198268, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
```

```

; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 198268
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-198268

Query Match
Best Local Similarity 100.0%; Pred. No. 2,2e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCCTGGCCAAACAT 2940
DB 322 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCCTGGCCAAACAT 367

RESULT 62
US-09-949-016-198269
; Sequence 198269, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 198269
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-198269

Query Match
Best Local Similarity 100.0%; Pred. No. 2,2e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCCTGGCCAAACAT 2940
DB 250 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCCTGGCCAAACAT 295

RESULT 63
US-09-949-016-198270
; Sequence 198270, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 198270
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-198270

Query Match
Best Local Similarity 100.0%; Pred. No. 2,2e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCCTGGCCAAACAT 2940
DB 60 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCCTGGCCAAACAT 105

RESULT 64
US-09-949-016-206260
; Sequence 206260, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 206260
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-206260

Query Match
Best Local Similarity 100.0%; Pred. No. 2,2e-11;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCCTGGCCAAACAT 2940
DB 57 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCCTGGCCAAACAT 102

RESULT 65
US-09-949-016-206261
; Sequence 206261, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 206261
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
```

US-09-949-016-206261

Query Match 1.5% Score 46; DB 3; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2,2e-11;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2895 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 2940  
|||||  
Db 74 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 119

RESULT 66  
US-09-949-016-206262  
; Sequence 206262, Application US/09949016  
; Patent No. 6812339

GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 206262

LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-206262

Query Match 1.5% Score 46; DB 3; Length 601;  
Best Local Similarity 100.0%; Pred. No. 2,2e-11;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2895 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 2940  
|||||  
Db 90 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 135

RESULT 67  
US-09-949-016-14878  
; Sequence 14878, Application US/09949016  
; Patent No. 6812339

GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14878

LENGTH: 17704  
TYPE: DNA  
ORGANISM: Human

FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(17704)  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14878

Query Match 1.5% Score 46; DB 3; Length 17704;  
Best Local Similarity 100.0%; Pred. No. 2e-11;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2895 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 2940  
|||||  
Db 8184 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 8229

RESULT 68  
US-09-949-016-15290/c  
; Sequence 15290, Application US/09949016  
; Patent No. 6812339

GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15290

LENGTH: 27600  
TYPE: DNA  
ORGANISM: Human

FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(27600)  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-15290

Query Match 1.5% Score 46; DB 3; Length 27600;  
Best Local Similarity 100.0%; Pred. No. 2e-11;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2895 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 2940  
|||||  
Db 3783 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 3738

RESULT 69  
US-09-949-016-11978/c  
; Sequence 11978, Application US/09949016  
; Patent No. 6812339

GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 11978

LENGTH: 29960  
TYPE: DNA

ORGANISM: Human  
US-09-949-016-11978

Query Match 1.5%; Score 46; DB 3; Length 29960;  
Best Local Similarity 100.0%; Pred. No. 1.9e-11;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 2940  
DB 8547 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 8502

RESULT 70  
US-09-949-016-16100/c  
; Sequence 16100, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16100  
; LENGTH: 30054  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-16100

Query Match 1.5%; Score 46; DB 3; Length 30054;  
Best Local Similarity 100.0%; Pred. No. 1.9e-11;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 2940  
DB 8640 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 8595

RESULT 71  
US-09-949-016-16864/c  
; Sequence 16864, Application US/09949016  
; Patent No. 6812338  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16864  
; LENGTH: 40905  
; TYPE: DNA  
; ORGANISM: Human  
; NAME/KEY: misc. feature  
; LOCATION: (1)...(40905)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-16864

Query Match 1.5%; Score 46; DB 3; Length 40905;

Best Local Similarity 100.0%; Pred. No. 1.9e-11;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 2940  
DB 22092 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 22047

RESULT 72  
US-09-949-016-12638/c  
; Sequence 12638, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12638  
; LENGTH: 60917  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-12638

Query Match 1.5%; Score 46; DB 3; Length 60917;  
Best Local Similarity 100.0%; Pred. No. 1.9e-11;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 2940  
DB 39793 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 39748

RESULT 73  
US-09-949-016-16057/c  
; Sequence 16057, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16057  
; LENGTH: 60917  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-16057

Query Match 1.5%; Score 46; DB 3; Length 60917;  
Best Local Similarity 100.0%; Pred. No. 1.9e-11;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 2940

Db 39793 GGTGATCATCTGAGGCCAGAGTTTCGAGACCAAGCCTGCGCAACAT 39748

RESULT 74  
US-09-949-016-14736/c  
Sequence 14736, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 14736  
LENGTH: 84916  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(84916)  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-14736

Query Match 1.5%; Score 46; DB 3; Length 84916;  
Best Local Similarity 100.0%; Pred. No. 1.9e-11;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2895 GGTGATCATCTGAGGCCAGAGTTTCGAGACCAAGCCTGCGCAACAT 2940  
Db 15607 GGTGATCATCTGAGGCCAGAGTTTCGAGACCAAGCCTGCGCAACAT 15562

RESULT 75  
US-09-949-016-15606  
Sequence 15606, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 15606  
LENGTH: 96739  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(96739)  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-15606

Query Match 1.5%; Score 46; DB 3; Length 96739;  
Best Local Similarity 100.0%; Pred. No. 1.9e-11;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTCACCTGACCTCCAGGCTGGGCAAGAGAGACTC 3116  
Db 49090 CAAGATTGTCACCTGACCTCCAGGCTGGGCAAGAGAGACTC 49135

RESULT 76  
US-09-949-016-17565/c  
Sequence 17565, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 17565  
LENGTH: 116955  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-17565

Query Match 1.5%; Score 46; DB 3; Length 116955;  
Best Local Similarity 100.0%; Pred. No. 1.9e-11;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2895 GGTGATCATCTGAGGCCAGAGTTTCGAGACCAAGCCTGCGCAACAT 2940  
Db 52491 GGTGATCATCTGAGGCCAGAGTTTCGAGACCAAGCCTGCGCAACAT 52446

RESULT 77  
US-09-949-016-12439/c  
Sequence 12439, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12439  
LENGTH: 116136  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(116136)  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12439

Query Match 1.5%; Score 46; DB 3; Length 116136;  
Best Local Similarity 100.0%; Pred. No. 1.9e-11;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy	2895	GGTGGATCACTGAGGCCAGAGTTTCGAGACCAAGCCTGGCCAA	CAT	2940
Db	15502	GGTGGATCACTGAGGCCAGAGTTCGAGACCAAGCCTGGCCAA	CAT	15457

```

RESULT 78
US-09-949-016-15316/c
/ Sequence 15316, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 15316
/ LENGTH: 144322
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc feature
/ LOCATION: (1).._(144322)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15316

```

```

RESULT 79
US-09-949-016-16513/C
; Sequence 16513, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16513
; LENGTH: 177293
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16513

```

Db 34975 CAAGATTGTGCTGCCTCCAGCTGGGCA CAGAGCAAGACTC 34930

```

RESULT 80
US-09-949-016-16264
; Sequence 16264, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16264
; LENGTH: 240157
; TYPE: DNA
; ORGANISM: Human
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)...(240157)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16264

```

```

RESULT 81
US-09-949-016-12656/C
; Sequence 12656, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12656
; LENGTH: 253345
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12656

```

Db 62618 GGTGATGACCTGAGGCCAGAGTTGAGACCAAGCCTGGCCAAAT 62573

## RESULT 82

US-09-949-016-13639/c  
; Sequence 13639, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13639  
; LENGTH: 253364  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-13639

Query Match 1.5%; Score 46; DB 3; Length 253364;  
Best Local Similarity 100.0%; Pred. No. 1.8e-11;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATGACCTGAGGCCAGAGTTGAGACCAAGCCTGGCCAAAT 2940  
DB 62618 GGTGATGACCTGAGGCCAGAGTTGAGACCAAGCCTGGCCAAAT 62573

## RESULT 83

US-09-949-016-12008/c  
; Sequence 12008, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12008  
; LENGTH: 331814  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(331814)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12008

Query Match 1.5%; Score 46; DB 3; Length 331814;  
Best Local Similarity 100.0%; Pred. No. 1.8e-11;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTC 3116  
DB 253007 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTC 252962

## RESULT 84

US-09-949-016-17056/c  
; Sequence 17056, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17056  
; LENGTH: 331814  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(331814)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-17056

Query Match 1.5%; Score 46; DB 3; Length 331814;  
Best Local Similarity 100.0%; Pred. No. 1.8e-11;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTC 3116  
DB 253007 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTC 252962

## RESULT 85

US-09-949-016-12147/c  
; Sequence 12147, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12147  
; LENGTH: 767677  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(767677)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-12147

Query Match 1.5%; Score 46; DB 3; Length 767677;  
Best Local Similarity 100.0%; Pred. No. 1.8e-11;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;





```
; Sequence 30239, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30239
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-30239

Query Match      1.4%; Score 45; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 6.5e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2898 GGATCCTGAGGCGGAGGAGTTGAGACCAAGCTGCGCAACATAG 2942
DB      328 GGATCCTGAGGCGGAGGAGTTGAGACCAAGCTGCGCAACATAG 372
|||||
RESULT 91
US-09-949-016-65461/c
; Sequence 65461, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 65461
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-65461

Query Match      1.4%; Score 45; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 6.5e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3078 GTGGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122
DB      577 GTGGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 533
|||||
RESULT 92
US-09-949-016-72556
; Sequence 72556, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72556
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-72556

Query Match      1.4%; Score 45; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 6.5e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3078 GTGGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122
DB      71 GTGGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 115
|||||
RESULT 93
US-09-949-016-72557
; Sequence 72557, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-20
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-10-03
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 72557
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-72557

Query Match      1.4%; Score 45; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 6.5e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3078 GTGGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122
DB      529 GTGGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 573
|||||
RESULT 94
US-09-949-016-82095
; Sequence 82095, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
```

```

; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 82095
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-82095

```

```

Query Match          1.4%; Score 45; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 6.5e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2898 GGATCAGCTGAGGCCGAGGAGTTGAGACCAAGCTGCGCAACATAG 2942
DB      535 GGATCAGCTGAGGCCGAGGAGTTGAGACCAAGCTGCGCAACATAG 579

```

```

RESULT 95
US-09-949-016-82096
; Sequence 82096, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 82096
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-82096

```

```

Query Match          1.4%; Score 45; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 6.5e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2898 GGATCAGCTGAGGCCGAGGAGTTGAGACCAAGCTGCGCAACATAG 2942
DB      401 GGATCAGCTGAGGCCGAGGAGTTGAGACCAAGCTGCGCAACATAG 445

```

```

RESULT 96
US-09-949-016-82097
; Sequence 82097, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08

```

```

; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 82097
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-82097

```

```

Query Match          1.4%; Score 45; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 6.5e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      2898 GGATCAGCTGAGGCCGAGGAGTTGAGACCAAGCTGCGCAACATAG 2942
DB      328 GGATCAGCTGAGGCCGAGGAGTTGAGACCAAGCTGCGCAACATAG 372

```

```

RESULT 97
US-09-949-016-145830/c
; Sequence 145830, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 145830
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-145830

```

```

Query Match          1.4%; Score 45; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 6.5e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3078 GTGCACCTGCACTCCAGCTGCGCAACAGCAAGACTGTCTC 3122
DB      502 GTGCACCTGCACTCCAGCTGCGCAACAGCAAGACTGTCTC 458

```

```

RESULT 98
US-09-949-016-146098/c
; Sequence 146098, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 146098
; LENGTH: 601
; TYPE: DNA

```

ORGANISM: Human  
US-09-949-016-146098

Query Match 1.4%; Score 45; DB 3; Length 601;  
Best Local Similarity 100.0%; Pred. No. 6.5e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTCTGTCTC 3122  
DB 502 GTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTCTGTCTC 458

RESULT 99  
US-09-949-016-146366/c  
Sequence 146366, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 146366  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-146366

Query Match 1.4%; Score 45; DB 3; Length 601;  
Best Local Similarity 100.0%; Pred. No. 6.5e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTCTGTCTC 3122  
DB 502 GTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTCTGTCTC 458

RESULT 100  
US-09-949-016-166442  
Sequence 166442, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 166442  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-166442

Query Match 1.4%; Score 45; DB 3; Length 601;  
Best Local Similarity 100.0%; Pred. No. 6.5e-11;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCACTGCACTCCAGCCTGGGCAACAGACAAGACT 3115  
DB 377 CAAGATTGTGCACTGCACTCCAGCCTGGGCAACAGACAAGACT 421

RESULT 101  
US-09-949-016-169130/c  
Sequence 169130, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 169130  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-169130

Query Match 1.4%; Score 45; DB 3; Length 601;  
Best Local Similarity 100.0%; Pred. No. 6.5e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCACTGCACTCCAGCCTGGGCAACAGACAAGACT 3115  
DB 519 CAAGATTGTGCACTGCACTCCAGCCTGGGCAACAGACAAGACT 475

RESULT 102  
US-09-949-016-169131/c  
Sequence 169131, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 169131  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-169131

Query Match 1.4%; Score 45; DB 3; Length 601;  
Best Local Similarity 100.0%; Pred. No. 6.5e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCACTGCACTCCAGCCTGGGCAACAGACAAGACT 3115  
DB 236 CAAGATTGTGCACTGCACTCCAGCCTGGGCAACAGACAAGACT 192

```
RESULT 103
US-09-949-016-203717
; Sequence 203717, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 203717
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-203717

Query Match          1.4%; Score 45; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 6.5e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2898 GGATCAGCTGAGGCGCAGAGTTGAGACCAAGCCTGGCCAAACATAG 2942
DB      535 GGATCAGCTGAGGCGCAGAGTTGAGACCAAGCCTGGCCAAACATAG 579

RESULT 104
US-09-949-016-203718
; Sequence 203718, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 203718
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-203718

Query Match          1.4%; Score 45; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 6.5e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2898 GGATCAGCTGAGGCGCAGAGTTGAGACCAAGCCTGGCCAAACATAG 2942
DB      401 GGATCAGCTGAGGCGCAGAGTTGAGACCAAGCCTGGCCAAACATAG 445

RESULT 105
US-09-949-016-203719
; Sequence 203719, Application US/09949016
```

```
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 203719
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-203719

Query Match          1.4%; Score 45; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 6.5e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2898 GGATCAGCTGAGGCGCAGAGTTGAGACCAAGCCTGGCCAAACATAG 2942
DB      328 GGATCAGCTGAGGCGCAGAGTTGAGACCAAGCCTGGCCAAACATAG 372

RESULT 106
US-09-949-002-4958
; Sequence 4958, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4958
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-4958

Query Match          1.4%; Score 45; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 6.5e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3078 GTGCCACTGCACTCGAGCTGGGCAACAGACCAAGCTCTGCTTC 3122
DB      524 GTGCCACTGCACTCGAGCTGGGCAACAGACCAAGCTCTGCTTC 568

RESULT 107
US-09-949-002-4959
; Sequence 4959, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
```

```

; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4959
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-4959
```

```

Query Match
Best Local Similarity 1.4%; Score 45; DB 3; Length 601;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGACGAACTGTCTC 3122
DB 521 GTGCCACTGCACTCCAGCCTGGGCAACAGACGAACTGTCTC 565
```

```

RESULT 108
US-09-949-002-9370/c
; Sequence 9370, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9370
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-9370
```

```

Query Match
Best Local Similarity 1.4%; Score 45; DB 3; Length 601;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGACGAACTGTCTC 3122
DB 489 GTGCCACTGCACTCCAGCCTGGGCAACAGACGAACTGTCTC 445
```

```

RESULT 109
US-09-949-016-14483/c
; Sequence 14483, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14483
; LENGTH: 8429
; TYPE: DNA
```

```

; ORGANISM: Human
US-09-949-016-14483
```

```

Query Match
Best Local Similarity 1.4%; Score 45; DB 3; Length 8429;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3073 AGATTGGCACTGCACTCCAGCCTGGGCAACAGACGAACTCT 3117
DB 4310 AGATTGGCACTGCACTCCAGCCTGGGCAACAGACGAACTCT 4266
```

```

RESULT 110
US-09-949-016-14852/c
; Sequence 14852, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 14852
; LENGTH: 9123
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-14852
```

```

Query Match
Best Local Similarity 1.4%; Score 45; DB 3; Length 9123;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGACGAACTGTCTC 3122
DB 664 GTGCCACTGCACTCCAGCCTGGGCAACAGACGAACTGTCTC 620
```

```

RESULT 111
US-09-949-016-17432/c
; Sequence 17432, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
```

```

; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17432
; LENGTH: 9176
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17432
```

```

Query Match
Best Local Similarity 1.4%; Score 45; DB 3; Length 9176;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy	3078 GTGCCACTGCACCTCCAGCCTGGGCAACAGAGCAAGACTGTGTC	3122
D6	664 GTGCCACTGCACCTCCAGCCTGGGCAACAGAGCAAGACTGTGTC	620

RESULT 112  
 US-09-949-016-13766  
 Sequence 13766, Application US/09949016  
 Patent No. 6812319  
 GENERAL INFORMATION:  
 APPLICANT: VENTER, J. Craig et al.  
 TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 FILE REFERENCE: C1001307  
 CURRENT APPLICATION NUMBER: US/09/949, 016  
 CURRENT FILING DATE: 2000-04-14  
 PRIOR APPLICATION NUMBER: 60/241,755  
 PRIOR FILING DATE: 2000-10-20  
 PRIOR APPLICATION NUMBER: 60/237,768  
 PRIOR FILING DATE: 2000-10-03  
 PRIOR APPLICATION NUMBER: 60/231,498  
 PRIOR FILING DATE: 2000-09-08  
 NUMBER OF SEQ ID NOS: 207012  
 SOFTWARE: fastseq for Windows Version 4.0  
 SEQ ID NO 13766  
 LENGTH: 10597  
 TYPE: DNA  
 ORGANISM: Human  
 US-09-949-016-13766

Query Match 1.4%; Score 45; DB 3; Length 10597;  
 Best Local Similarity 100.0%; Pred. No. 5.9e-11;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy		3078 GTGCCACTGCACCTCCAGCCTGGGCAACAGAGCAAGA CTGTGTCTC	3122
Dd		952 GTGCCA CTGCAC TCCAG CCGTGG CAACA GAGCA AGA CTGTGTCTC	996

RESULT 113  
 US-09-949-016-17469/C  
 / Sequence 17469, Application US/09949016  
 / Patent No. 6812339  
 / GENERAL INFORMATION:  
 / APPLICANT: VENTER, J. Craig et al.  
 / TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 / TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 / FILE REFERENCE: CLO01107  
 / CURRENT APPLICATION NUMBER: US/09/949,016  
 / CURRENT FILING DATE: 2000-04-14  
 / PRIOR APPLICATION NUMBER: 60/241,755  
 / PRIOR FILING DATE: 2000-10-20  
 / PRIOR APPLICATION NUMBER: 60/237,768  
 / PRIOR FILING DATE: 2000-10-03  
 / PRIOR APPLICATION NUMBER: 60/231,498  
 / PRIOR FILING DATE: 2000-09-08  
 / NUMBER OF SEQ ID NOS: 207012  
 / SOFTWARE: FASTSEQ for Windows Version 4.0  
 / SEQ ID NO 17469  
 / LENGTH: 15975  
 / TYPE: DNA  
 / ORGANISM: Human  
 / US-09-949-016-17469

Query Match	1.4%;	Score 45;	DB 3;	Length 15975;
Best Local Similarity	100.0%;	Pred. No. 5.8e-11;		
Matches	45;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0;

QY	3078	GTGCACTGCACTCCAGCCTGGCAACAGAGCAAGACTCTGTCTC	3122
Db	8586	GTGCCACTGCACTCCAGCCTGGCAACAGAGCAAGACTCTGTCTC	85422

RESULT 114  
US-09-146-053-7  
; Sequence 7, Application US/09146053A

Query Match 1.4%; Score 45; DB 3; Length 16595;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-11;  
 Matches 45; Conservative 0; Mismatches 0; Gaps 0;

OY	3078 GTGCCACTGCACTCCAGCCTGGACAACAGAGAAGTCTGTCTC	3122
Dδ	4301 GTGCCACTGCACTCCAGCCTGGACAACAGAGAAGTCTGTCTC	4345

```

RESULT 115
US-09-949-016-16393/C
/ Sequence 16393, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE. METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CLO01307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO 16393
/ LENGTH: 19296
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-16393

```

Query Match 1.4%; Score 45; DB 3; Length 19296;  
 Best Local Similarity 100.0%; Pred. No. 5.8e-11;  
 Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	3078	GTGGCACTGTCACTCCAGCCTGGGCAACAGAGCAGACTCTGTCTC	3122
Db	14177	GTGGCACTGTCACTCCAGCCTGGGCAACAGAGCAGACTCTGTCTC	14133

RESULT 116  
US-09-949-016-15920/c  
; Sequence 15920, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.



```
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL001307
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 15920
/ LENGTH: 19943
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-15920
```

```
Query Match 1.4%; Score 45; DB 3; Length 19943;
Best Local Similarity 100.0%; Pred. No. 5.8e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2896 GTGGATCACTGAGGCCAGAGTTTCGAGACCAAGCTTGCCCAACAT 2940
Db 18806 GTGGATCACTGAGGCCAGAGTTTCGAGACCAAGCTTGCCCAACAT 18762
```

```
RESULT 117
US-09-949-016-12553/c
/ Sequence 12553, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 12553
/ LENGTH: 19945
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-12553
```

```
Query Match 1.4%; Score 45; DB 3; Length 19945;
Best Local Similarity 100.0%; Pred. No. 5.8e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2896 GTGGATCACTGAGGCCAGAGTTTCGAGACCAAGCTTGCCCAACAT 2940
Db 18806 GTGGATCACTGAGGCCAGAGTTTCGAGACCAAGCTTGCCCAACAT 18762
```

```
RESULT 118
US-09-949-016-13069/c
/ Sequence 13069, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
```

```
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 13069
/ LENGTH: 27545
/ TYPE: DNA
/ ORGANISM: Human
US-09-949-016-13069
```

```
Query Match 1.4%; Score 45; DB 3; Length 27545;
Best Local Similarity 100.0%; Pred. No. 5.7e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 3078 GTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122
Db 25807 GTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 25763
```

```
RESULT 119
US-09-949-016-16094
/ Sequence 16094, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO: 16094
/ LENGTH: 33099
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc.feature
/ LOCATION: (1)...(33099)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16094
```

```
Query Match 1.4%; Score 45; DB 3; Length 33099;
Best Local Similarity 100.0%; Pred. No. 5.7e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 3078 GTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122
Db 32755 GTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 32799
```

```
RESULT 120
US-09-949-016-14417/c
/ Sequence 14417, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ PRIOR FILING DATE: 2000-04-14
```



US-09-949-016-14128

Query Match 1.4%; Score 45; DB 3; Length 40379;  
Best Local Similarity 100.0%; Pred. No. 5.7e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2898 GGATCAGCTGAGCCAGAGGTTGAGACCAAGCTGGCCAAACATAG 2942

Db 26306 GGATCAGCTGAGCCAGAGGTTGAGACCAAGCTGGCCAAACATAG 26350

RESULT 125

US-09-949-016-17495  
Sequence 17495, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 17495  
LENGTH: 40877  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-17495

Query Match 1.4%; Score 45; DB 3; Length 40877;  
Best Local Similarity 100.0%; Pred. No. 5.7e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2898 GGATCAGCTGAGCCAGAGGTTGAGACCAAGCTGGCCAAACATAG 2942

Db 25815 GGATCAGCTGAGCCAGAGGTTGAGACCAAGCTGGCCAAACATAG 25859

RESULT 126

US-09-949-002-845  
Sequence 845, Application US/09949002  
Patent No. 6900016

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION  
AND USES THEREOF  
FILE REFERENCE: CL000790  
CURRENT APPLICATION NUMBER: US/09/949,002  
CURRENT FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: 60/231,401  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 10823  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 845  
LENGTH: 40908  
TYPE: DNA  
ORGANISM: Human  
US-09-949-002-845

Query Match 1.4%; Score 45; DB 3; Length 40908;  
Best Local Similarity 100.0%; Pred. No. 5.7e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2906 TGAGCCAGAGTTGAGACCAAGCTGGCCAAACATAGCAAAACC 2950

|||||

Db 16225 TGAGCCAGAGTTGAGACCAAGCTGGCCAAACATAGCAAAACC 16269

RESULT 127  
US-09-949-002-603  
Sequence 603, Application US/09949002  
Patent No. 6900016

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION  
AND USES THEREOF  
FILE REFERENCE: CL000790  
CURRENT APPLICATION NUMBER: US/09/949,002  
CURRENT FILING DATE: 2000-01-28  
PRIOR APPLICATION NUMBER: 60/231,401  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 10823  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 603  
LENGTH: 40944  
TYPE: DNA  
ORGANISM: Human  
US-09-949-002-603

Query Match 1.4%; Score 45; DB 3; Length 40944;  
Best Local Similarity 100.0%; Pred. No. 5.7e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2906 TGAGCCAGAGTTGAGACCAAGCTGGCCAAACATAGCAAAACC 2950

Db 16262 TGAGCCAGAGTTGAGACCAAGCTGGCCAAACATAGCAAAACC 16306

RESULT 128

US-09-949-016-16492  
Sequence 16492, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 16492  
LENGTH: 47781  
TYPE: DNA  
ORGANISM: Human  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (1)...(4781)  
OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-16492

Query Match 1.4%; Score 45; DB 3; Length 47781;  
Best Local Similarity 100.0%; Pred. No. 5.6e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 GTGACCTGAGCTGAGCTGGGCAAGAGCAAGACTGTGTCTC 3122

Db 46796 GTGACCTGAGCTGAGCTGGGCAAGAGCAAGACTGTGTCTC 46840

RESULT 129

|||||

```
US-09-949-016-16493
; Sequence 16493, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16493
; LENGTH: 47781
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(47781)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16493
```

```
Query Match 1.4%; Score 45; DB 3; Length 47781;
Best Local Similarity 100.0%; Pred. No. 5.6e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3078 GTGCCACTGCATCTCCAGCTGGGCAACAGACAAGACTCTGTCTC 3122
DB 46796 GTGCCACTGCATCTCCAGCTGGGCAACAGACAAGACTCTGTCTC 46840
```

```
RESULT 130
US-09-949-016-16494
; Sequence 16494, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16494
; LENGTH: 47781
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(47781)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16494
```

```
Query Match 1.4%; Score 45; DB 3; Length 47781;
Best Local Similarity 100.0%; Pred. No. 5.6e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3078 GTGCCACTGCATCTCCAGCTGGGCAACAGACAAGACTCTGTCTC 3122
DB 46796 GTGCCACTGCATCTCCAGCTGGGCAACAGACAAGACTCTGTCTC 46840
```

```
RESULT 131
US-09-949-002-817/c
; Sequence 817, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; AND USES THEREOF
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 817
; LENGTH: 49677
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-817
```

```
Query Match 1.4%; Score 45; DB 3; Length 49677;
Best Local Similarity 100.0%; Pred. No. 5.6e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3078 GTGCCACTGCATCTCCAGCTGGGCAACAGACAAGACTCTGTCTC 3122
DB 19485 GTGCCACTGCATCTCCAGCTGGGCAACAGACAAGACTCTGTCTC 19441
```

```
RESULT 132
US-09-949-016-15254
; Sequence 15254, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15254
; LENGTH: 64984
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(64984)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15254
```

```
Query Match 1.4%; Score 45; DB 3; Length 64984;
Best Local Similarity 100.0%; Pred. No. 5.6e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3078 GTGCCACTGCATCTCCAGCTGGGCAACAGACAAGACTCTGTCTC 3122
DB 35279 GTGCCACTGCATCTCCAGCTGGGCAACAGACAAGACTCTGTCTC 35323
```

```
RESULT 133
US-09-949-016-12630/c
```

```
; Sequence 12630, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12630
; LENGTH: 66627
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12630

Query Match      1.4%; Score 45; DB 3; Length 66627;
Best Local Similarity 100.0%; Pred. No. 5.6e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3078 GTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTCTGTCTC 3122
DB      12900 GTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTCTGTCTC 12856

RESULT 134
US-09-949-016-16112/c
; Sequence 16112, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16112
; LENGTH: 66628
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16112

Query Match      1.4%; Score 45; DB 3; Length 66628;
Best Local Similarity 100.0%; Pred. No. 5.6e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3078 GTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTCTGTCTC 3122
DB      12900 GTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTCTGTCTC 12856

RESULT 135
US-09-949-002-714
; Sequence 714, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
```

```
; TITLE OF INVENTION: WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 714
; LENGTH: 70313
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-714

Query Match      1.4%; Score 45; DB 3; Length 70313;
Best Local Similarity 100.0%; Pred. No. 5.6e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3078 GTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTCTGTCTC 3122
DB      37513 GTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTCTGTCTC 37557

RESULT 136
US-09-949-002-633
; Sequence 633, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; PRIOR FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 633
; LENGTH: 72347
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-633

Query Match      1.4%; Score 45; DB 3; Length 72347;
Best Local Similarity 100.0%; Pred. No. 5.6e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3078 GTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTCTGTCTC 3122
DB      37513 GTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTCTGTCTC 37557

RESULT 137
US-09-949-016-12659/c
; Sequence 12659, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; NUMBER OF SEQ ID NOS: 207012
```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12659
; LENGTH: 80858
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-12659

Query Match          1.4%; Score 45; DB 3; Length 80858;
Best Local Similarity 100.0%; Pred. No. 5.5e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3078 GTGCCACTGCACTCCAGCTGGGCAACAGACAAGCAACTCTGTCTC 3122
Db      11289 GTGCCACTGCACTCCAGCTGGGCAACAGACAAGCAACTCTGTCTC 11245

RESULT 138
US-09-949-016-15715/c
; Sequence 15715, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15715
; LENGTH: 80859
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15715

Query Match          1.4%; Score 45; DB 3; Length 80859;
Best Local Similarity 100.0%; Pred. No. 5.5e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3078 GTGCCACTGCACTCCAGCTGGGCAACAGACAAGCAACTCTGTCTC 3122
Db      11289 GTGCCACTGCACTCCAGCTGGGCAACAGACAAGCAACTCTGTCTC 11245

RESULT 139
US-09-949-016-13655
; Sequence 13655, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13655
; LENGTH: 87648
; TYPE: DNA
; ORGANISM: Human

; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(87648)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-13655

Query Match          1.4%; Score 45; DB 3; Length 87648;
Best Local Similarity 100.0%; Pred. No. 5.5e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3078 GTGCCACTGCACTCCAGCTGGGCAACAGACAAGCAACTCTGTCTC 3122
Db      35917 GTGCCACTGCACTCCAGCTGGGCAACAGACAAGCAACTCTGTCTC 35961

RESULT 140
US-09-949-016-15146
; Sequence 15146, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15146
; LENGTH: 91279
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15146

Query Match          1.4%; Score 45; DB 3; Length 91279;
Best Local Similarity 100.0%; Pred. No. 5.5e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3078 GTGCCACTGCACTCCAGCTGGGCAACAGACAAGCAACTCTGTCTC 3122
Db      68268 GTGCCACTGCACTCCAGCTGGGCAACAGACAAGCAACTCTGTCTC 68312

RESULT 141
US-09-949-016-15944/c
; Sequence 15944, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15944
; LENGTH: 93532
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-15944
```

Query Match 1.4%; Score 45; DB 3; Length 93532;  
Best Local Similarity 100.0%; Pred. No. 5.5e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3078 GTGCCACTGCACCTCCAGCTGGGCAACAGACGAACTCTGTCTC 3122  
DB 20699 GTGCCACTGCACCTCCAGCTGGGCAACAGACGAACTCTGTCTC 20655

RESULT 142  
US-09-949-016-16553  
; Sequence 16553, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16553  
; LENGTH: 94142  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURES:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(94142)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-16553

Query Match 1.4%; Score 45; DB 3; Length 94142;  
Best Local Similarity 100.0%; Pred. No. 5.5e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3071 CAAGTTGTGCACCTCCAGCTGGGCAACAGACGAACTCT 3115  
DB 19351 CAAGTTGTGCACCTCCAGCTGGGCAACAGACGAACTCT 19395

RESULT 143  
US-09-949-016-17103  
; Sequence 17103, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17103  
; LENGTH: 96690  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17103

Query Match 1.4%; Score 45; DB 3; Length 96690;  
Best Local Similarity 100.0%; Pred. No. 5.5e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3078 GTGCCACTGCACCTCCAGCTGGGCAACAGACGAACTCTGTCTC 3122  
DB 72113 GTGCCACTGCACCTCCAGCTGGGCAACAGACGAACTCTGTCTC 72157

RESULT 144  
US-09-949-016-13319/c  
; Sequence 13319, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13319  
; LENGTH: 103750  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-13319

Query Match 1.4%; Score 45; DB 3; Length 103750;  
Best Local Similarity 100.0%; Pred. No. 5.5e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACGAACTCT 3117  
DB 75631 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACGAACTCT 75587

RESULT 145  
US-09-949-016-16424  
; Sequence 16424, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CLO01307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 16424  
; LENGTH: 112465  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-16424

Query Match 1.4%; Score 45; DB 3; Length 112465;  
Best Local Similarity 100.0%; Pred. No. 5.5e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3071 CAAGTTGTGCACCTCCAGCTGGGCAACAGACGAACTCT 3115



Db 35617 CAAGATTGCGCACTGCACTCCAGCTGGGCAACAGAGCAAGACT 35661

RESULT 146  
US-09-949-016-13915  
; Sequence 13915, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13915  
; LENGTH: 120609  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(120609)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-13915

Query Match 1.4%; Score 45; DB 3; Length 120609;  
Best Local Similarity 100.0%; Pred. No. 5.5e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTGCTC 3122  
Db 83825 GTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTGCTC 83869

RESULT 147  
US-09-949-016-17074/c  
; Sequence 17074, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17074  
; LENGTH: 140725  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-17074

Query Match 1.4%; Score 45; DB 3; Length 140725;  
Best Local Similarity 100.0%; Pred. No. 5.4e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTGCTC 3122  
|||||

Db 79776 GTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTGCTC 79732

RESULT 148  
US-09-949-016-13870/c  
; Sequence 13870, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13870  
; LENGTH: 155266  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-13870

Query Match 1.4%; Score 45; DB 3; Length 155266;  
Best Local Similarity 100.0%; Pred. No. 5.4e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTGCTC 3122  
Db 16826 GTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTGCTC 16782

RESULT 149  
US-09-949-016-13733  
; Sequence 13733, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13733  
; LENGTH: 162841  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-13733

Query Match 1.4%; Score 45; DB 3; Length 162841;  
Best Local Similarity 100.0%; Pred. No. 5.4e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGCGCACTGCACTCCAGCTGGGCAACAGAGCAAGACT 3115  
Db 119589 CAAGATTGCGCACTGCACTCCAGCTGGGCAACAGAGCAAGACT 119633

RESULT 150  
US-09-949-016-16509/c

```
/ Sequence 16509, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14, 755
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 16509
/ LENGTH: 174639
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(174639)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16509
```

```
Query Match 1.4%; Score 45; DB 3; Length 174639;
Best Local Similarity 100.0%; Pred. No. 5.4e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 3071 GAGAGTTGGCAGCTCCAGCTGGGCAACAGAGCAAGACT 3115
Db 107661 CAGAGTTGGCAGCTCCAGCTGGGCAACAGAGCAAGACT 107617
```

```
RESULT 151
US-09-949-016-15851
/ Sequence 15851, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 15851
/ LENGTH: 205044
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(205044)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15851
```

```
Query Match 1.4%; Score 45; DB 3; Length 205044;
Best Local Similarity 100.0%; Pred. No. 5.4e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 3078 GTGCCACTGCATCTCCAGCTGGGCAACAGAGCAAGACTGTGCTC 3122
Db 161139 GTGCCACTGCATCTCCAGCTGGGCAACAGAGCAAGACTGTGCTC 161183
```

```
RESULT 152
US-09-949-016-15852
/ Sequence 15852, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 15852
/ LENGTH: 205044
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(205044)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15852
```

```
Query Match 1.4%; Score 45; DB 3; Length 205044;
Best Local Similarity 100.0%; Pred. No. 5.4e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 3078 GTGCCACTGCATCTCCAGCTGGGCAACAGAGCAAGACTGTGCTC 3122
Db 161139 GTGCCACTGCATCTCCAGCTGGGCAACAGAGCAAGACTGTGCTC 161183
```

```
RESULT 153
US-09-949-016-15853
/ Sequence 15853, Application US/09949016
/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 15853
/ LENGTH: 205044
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(205044)
/ OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15853
```

```
Query Match 1.4%; Score 45; DB 3; Length 205044;
Best Local Similarity 100.0%; Pred. No. 5.4e-11;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 3078 GTGCCACTGCATCTCCAGCTGGGCAACAGAGCAAGACTGTGCTC 3122
```

Db 161139 GTGGCAGTGCATCCAGCCTGGGCAAGAGCAAGACTCTGTCTC 161183

```
RESULT 154
US-09-949-016-17539/c
; Sequence 17539, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17539
; LENGTH: 218940
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-17539
```

Query Match 1.4%; Score 45; DB 3; Length 218940;  
Best Local Similarity 100.0%; Pred.No. 5.4e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2978 GCCGGGCGATGTGGGCGGATGCTGTGCTCCAGCTACTCGGAGG 3022  
Db 109636 GCCGGGCGATGTGGGCGGATGCTGTGCTCCAGCTACTCGGAGG 109592

```
RESULT 155
US-09-949-016-12387
; Sequence 12387, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12387
; LENGTH: 223471
; TYPE: DNA
; ORGANISM: Human
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)...(223471)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12387
```

Query Match 1.4%; Score 45; DB 3; Length 223471;  
Best Local Similarity 100.0%; Pred.No. 5.4e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 GTGGCAGTGCATCCAGCCTGGGCAAGAGCAAGACTCTGTCTC 3122  
|||||

Db 131566 GTGGCAGTGCATCCAGCCTGGGCAAGAGCAAGACTCTGTCTC 131610

```
RESULT 156
US-09-949-016-12724
; Sequence 12724, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12724
; LENGTH: 223471
; TYPE: DNA
; ORGANISM: Human
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)...(223471)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12724
```

Query Match 1.4%; Score 45; DB 3; Length 223471;  
Best Local Similarity 100.0%; Pred.No. 5.4e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 GTGGCAGTGCATCCAGCCTGGGCAAGAGCAAGACTCTGTCTC 3122  
Db 131566 GTGGCAGTGCATCCAGCCTGGGCAAGAGCAAGACTCTGTCTC 131610  
|||||

```
RESULT 157
US-09-949-016-12725
; Sequence 12725, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12725
; LENGTH: 223471
; TYPE: DNA
; ORGANISM: Human
; FEATURES:
; NAME/KEY: misc_feature
; LOCATION: (1)...(223471)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12725
```

Query Match 1.4%; Score 45; DB 3; Length 223471;  
Best Local Similarity 100.0%; Pred.No. 5.4e-11;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 3122  
DB 131566 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 131610

## RESULT 158

US-09-513-999C-25910/c  
Sequence 25910, Application US/09513999C  
Patent No. 6783961

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Duclerc, A.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

Patent No. 6783961

FILE REFERENCE: 59. US2, REG

CURRENT APPLICATION NUMBER: US/09/513,999C

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/122,487

PRIOR FILING DATE: 1999-02-26

NUMBER OF SEQ ID NOS: 36681

SOFTWARE: Patent.pm

SEQ ID NO 25910

LENGTH: 145

TYPE: DNA

ORGANISM: Homo sapiens

FEATURE: misc\_feature

LOCATION: 120

OTHER INFORMATION: k=g or t

FEATURE: NAME/KEY: misc\_feature

LOCATION: 121

OTHER INFORMATION: r=a or g

US-09-513-999C-25910

Query Match 1.4%; Score 44; DB 3; Length 145;  
Best Local Similarity 100.0%; Pred. No. 2e-10;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 3122  
DB 47 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 4

## RESULT 159

US-09-513-999C-28305  
Sequence 28305, Application US/09513999C  
Patent No. 6783961

GENERAL INFORMATION:

APPLICANT: Dumas Milne Edwards, J.B.

APPLICANT: Duclerc, A.

APPLICANT: Giordano, J.Y.

TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.

Patent No. 6783961

FILE REFERENCE: 59. US2, REG

CURRENT APPLICATION NUMBER: US/09/513,999C

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/122,487

PRIOR FILING DATE: 1999-02-26

NUMBER OF SEQ ID NOS: 36681

SOFTWARE: Patent.pm

SEQ ID NO 28305

LENGTH: 330

TYPE: DNA

ORGANISM: Homo sapiens

US-09-513-999C-28305

Query Match 1.4%; Score 44; DB 3; Length 330;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2899 GATCACTGAGGCCAGAGTTGAGACCAAGCCTGGCCAATATG 2942  
DB 65 GATCACTGAGGCCAGAGTTGAGACCAAGCCTGGCCAATATG 108

## RESULT 160

US-09-949-016-33113  
Sequence 33113, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 33113

LENGTH: 601

TYPE: DNA

ORGANISM: Human

US-09-949-016-33113

Query Match 1.4%; Score 44; DB 3; Length 601;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3079 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 3122  
DB 430 TGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 473

## RESULT 161

US-09-949-016-36581  
Sequence 36581, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:

APPLICANT: VENTER, J. Craig et al.

TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED

WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF

FILE REFERENCE: CLO01307

CURRENT APPLICATION NUMBER: US/09/949,016

PRIOR FILING DATE: 2000-04-14

PRIOR APPLICATION NUMBER: 60/241,755

PRIOR FILING DATE: 2000-10-20

PRIOR APPLICATION NUMBER: 60/237,768

PRIOR FILING DATE: 2000-10-03

PRIOR APPLICATION NUMBER: 60/231,498

PRIOR FILING DATE: 2000-09-08

NUMBER OF SEQ ID NOS: 207012

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 36581

LENGTH: 601

TYPE: DNA

ORGANISM: Human

US-09-949-016-36581

Query Match 1.4%; Score 44; DB 3; Length 601;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;

Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 162
US-09-949-016-38261
; Sequence 38261, Application US/09949016
; Patent No 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: fastseq for Windows Version 4.0
; SEQ ID NO 38261
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-38261

```

```

Query Match          1.4%; Score 44; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

```
QY      3079 TGGCACTGCACTCCAGCCTGGGCAACAGAACAAGACTGTGTCTC 3122
          |||||
Db       430 TGCCA CTGC ACTT CA GCT GGGA CACA GAACA AGA CT GTGT CTC 473
```

RESULT 163  
 US-09-949-016-46100  
 ; Sequence 46100, Application US/09949016  
 ; Patent No. 6812339  
 ; GENERAL INFORMATION:  
 ; APPLICANT: VENTER, J. Craig et al.  
 ; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
 ; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
 ; FILE REFERENCE: C1001307  
 ; CURRENT APPLICATION NUMBER: US/09/949,016  
 ; CURRENT FILING DATE: 2000-04-14  
 ; PRIOR APPLICATION NUMBER: 60/241,755  
 ; PRIOR FILING DATE: 2000-10-20  
 ; PRIOR APPLICATION NUMBER: 60/237,768  
 ; PRIOR FILING DATE: 2000-10-03  
 ; PRIOR APPLICATION NUMBER: 60/231,498  
 ; PRIOR FILING DATE: 2000-09-08  
 ; NUMBER OF SEQ ID NOS: 207012  
 ; SOFTWARE: FASTSEQ for Windows Version 4.0  
 ; SEQ ID NO 46100  
 ; LENGTH: 601  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 ; US-09-949-016-46100

Query Match	1.4%;	Score 44;	DB 3;	Length 601;
Best Local Similarity	100.0%;	Pred. No. 1.9e-10;		
Matches 44;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

```

Oy      3073 AGATTGTGCCACTGCACTCCAGCCCTGGGCAACAGAGCAAGACTC 3116
          |||||
Db      154  AGATTGTGCCACTGCACTCCAGCCCTGGGCAACAGAGCAAGACTC 197

```

```

RESULT 164
US-09-949-016-46119
; Sequence 46119, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:

```

```

? APPLICANT: VENTER, J. Craig et al.
? TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
? TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
? FILE REFERENCE: C0001307
? CURRENT APPLICATION NUMBER: US/09/949,016
? CURRENT FILING DATE: 2000-04-14
? PRIOR APPLICATION NUMBER: 60/241,755
? PRIOR FILING DATE: 2000-10-20
? PRIOR APPLICATION NUMBER: 60/237,768
? PRIOR FILING DATE: 2000-10-03
? PRIOR APPLICATION NUMBER: 60/231,498
? PRIOR FILING DATE: 2000-09-08
? NUMBER OF SEQ ID NOS: 207012
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 46119
? LENGTH: 601
? TYPE: DNA
? ORGANISM: Human
? US-09-949-016-46119

```

Query Match	1.4%;	Score 44;	DB 3;	Length 601;
Best Local Similarity	100.0%;	Pred. No. 1.9e-10;		
Matches 44;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

QY      3073 AGATTGTGCCACTGCACCTCCAGCCTGGGCAACAGAGCAAGACTC 3116
          ||| | | | | | | | | | | | | | | | | | | | | | | | | |
DB      154 AGATTGTGCCACTGCACCTCCAGCCTGGGCAACAGAGCAAGACTC 197

```

```

RESULT 165
US-09-949-016-71064
: Sequence 71064, Application US/03949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CLO01107
: CURRENT APPLICATION NUMBER: US/09/949, 016
: CURRENT FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 71064
: LENGTH: 601
: TYPE: DNA
: ORGANISM: Human
US-09-949-016-71064

```

Query Match	1.4%	Score 44;	DB 3;	Length 601;
Best Local Similarity	100.0%	Pred. No. 1.9e-10;		
Matches 44;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

```

Oy      3073 AGATTGTGCCACTGCACCTCCAGCCTGGGCAACAGAGCAAGACTC 3116
          |||||
Db      59  AGATTGTGCCACTGCACCTCCAGCCTGGGCAACAGAGCAAGACTC 102

```

```

RESULT 166
US-09-949-016-71065
; Sequence 71065, Application US/09949016
; Patent No. 681239
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949, 016

```

```

; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 71065
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-71065

Query Match          1.4%; Score 44; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTGCGCACTGCACCTCGGCGAAGAGCAAGACTC 3116
Db      110 AGATTGTGCGCACTGCACCTCGGCGAAGAGCAAGACTC 153

RESULT 167
US-09-949-016-71066
; Sequence 71066, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 71066
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-71066

Query Match          1.4%; Score 44; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTGCGCACTGCACCTCGGCGAAGAGCAAGACTC 3116
Db      326 AGATTGTGCGCACTGCACCTCGGCGAAGAGCAAGACTC 369

RESULT 168
US-09-949-016-71067
; Sequence 71067, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
```

```

; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 71067
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-71067

Query Match          1.4%; Score 44; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTGCGCACTGCACCTCGGCGAAGAGCAAGACTC 3116
Db      359 AGATTGTGCGCACTGCACCTCGGCGAAGAGCAAGACTC 402

RESULT 169
US-09-949-016-71068
; Sequence 71068, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 71068
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-71068

Query Match          1.4%; Score 44; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTGCGCACTGCACCTCGGCGAAGAGCAAGACTC 3116
Db      379 AGATTGTGCGCACTGCACCTCGGCGAAGAGCAAGACTC 422

RESULT 170
US-09-949-016-71069
; Sequence 71069, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 71069
```

LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-71069

Query Match 1.4%; Score 44; DB 3; Length 601;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTCACCTGCACTCCAGCCTGGGCAACAGAGAAGACTC 3116  
DB 434 AGATTGTCACCTGCACTCCAGCCTGGGCAACAGAGAAGACTC 477

RESULT 171  
US-09-949-016-121487/c  
Sequence 121487, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
PRIOR FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 121487  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-121487

Query Match 1.4%; Score 44; DB 3; Length 601;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTCACCTGCACTCCAGCCTGGGCAACAGAGAAGACTC 3116  
DB 120 AGATTGTCACCTGCACTCCAGCCTGGGCAACAGAGAAGACTC 77

RESULT 172  
US-09-949-016-121488/c  
Sequence 121488, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
PRIOR FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 121488  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-121488

Query Match 1.4%; Score 44; DB 3; Length 601;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTCACCTGCACTCCAGCCTGGGCAACAGAGAAGACTC 3116  
DB 351 AGATTGTCACCTGCACTCCAGCCTGGGCAACAGAGAAGACTC 308

RESULT 173  
US-09-949-016-133933/c  
Sequence 133933, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
PRIOR FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 133933  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-133933

Query Match 1.4%; Score 44; DB 3; Length 601;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTCACCTGCACTCCAGCCTGGGCAACAGAGAAGACTC 3116  
DB 370 AGATTGTCACCTGCACTCCAGCCTGGGCAACAGAGAAGACTC 327

RESULT 174  
US-09-949-016-133934/c  
Sequence 133934, Application US/09949016  
Patent No. 6812339

GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
PRIOR FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 133934  
LENGTH: 601  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-133934

Query Match 1.4%; Score 44; DB 3; Length 601;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTCACCTGCACTCCAGCCTGGGCAACAGAGAAGACTC 3116



Db 368 AGATTGTGCACCTGCACCTCGGCAACAGACAGACTC 325

## RESULT 175

US-09-949-016-143264  
; Sequence 143264, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 143264  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-143264

Query Match 1.4%; Score 44; DB 3; Length 601;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCGGCAACAGACAGACTC 3116  
Db 233 AGATTGTGCACCTGCACCTCGGCAACAGACAGACTC 276

## RESULT 176

US-09-949-016-143435  
; Sequence 143435, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 143435  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-143435

Query Match 1.4%; Score 44; DB 3; Length 601;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCGGCAACAGACAGACTC 3116  
Db 233 AGATTGTGCACCTGCACCTCGGCAACAGACAGACTC 276

## RESULT 177

US-09-949-016-144443/c  
; Sequence 144443, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 144443  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-144443

Query Match 1.4%; Score 44; DB 3; Length 601;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TGCCACTGCACCTGCACCTCGGCAACAGACAGACTCTGTCTC 3122  
Db 158 TGCCACTGCACCTGCACCTCGGCAACAGACAGACTCTGTCTC 115

RESULT 178  
US-09-949-016-144444/c  
; Sequence 144444, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; PRIOR FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 144444  
; LENGTH: 601  
; TYPE: DNA  
; ORGANISM: Human  
US-09-949-016-144444

Query Match 1.4%; Score 44; DB 3; Length 601;  
Best Local Similarity 100.0%; Pred. No. 1.9e-10;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TGCCACTGCACCTGCACCTCGGCAACAGACAGACTCTGTCTC 3122  
Db 412 TGCCACTGCACCTGCACCTCGGCAACAGACAGACTCTGTCTC 369

RESULT 179  
US-09-949-016-150959  
; Sequence 150959, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.

```

; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 150959
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-150959

Query Match          1.4%; Score 44; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3073 AGATTGTGCGCACTGCACTCGAGCTGGGCAACAGAGCAAGACTC 3116
DB      233 AGATTGTGCGCACTGCACTCGAGCTGGGCAACAGAGCAAGACTC 276

RESULT 180
US-09-949-016-163060
; Sequence 163060, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 163060
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-163060

Query Match          1.4%; Score 44; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3079 TCCCACTGCACTCGAGCTGGGCAACAGAGCAAGACTCTGTC 3122
DB      158 TCCCACTGCACTCGAGCTGGGCAACAGAGCAAGACTCTGTC 201

RESULT 181
US-09-949-016-165548/c
; Sequence 165548, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
```

```

; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 165548
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-165548

Query Match          1.4%; Score 44; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3073 AGATTGTGCGCACTGCACTCGAGCTGGGCAACAGAGCAAGACTC 3116
DB      211 AGATTGTGCGCACTGCACTCGAGCTGGGCAACAGAGCAAGACTC 168

RESULT 182
US-09-949-016-199110/c
; Sequence 199110, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO: 199110
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-199110

Query Match          1.4%; Score 44; DB 3; Length 601;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      2889 GAGGAGGTGATCATCTGAGGCCAGGAGTTGAGACCAAGCTG 2932
DB      231 GAGGAGGTGATCATCTGAGGCCAGGAGTTGAGACCAAGCTG 188

RESULT 183
US-09-949-016-202729
; Sequence 202729, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
```

```

; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 202729
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-202729

Query Match
Best Local Similarity 100.0%; Score 44; DB 3; Length 601;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGCGCACTCCAGCTCGGCAACAGCAAGACTCTGCTC 3116
DB 231 AGATTGCGCACTCCAGCTCGGCAACAGCAAGACTCTGCTC 274

RESULT 184
US-09-949-016-204472
; Sequence 204472, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 204472
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-204472

Query Match
Best Local Similarity 100.0%; Score 44; DB 3; Length 601;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TCCCACTGCACTCCAGCTCGGCAACAGCAAGACTCTGCTC 3122
DB 430 TCCCACTGCACTCCAGCTCGGCAACAGCAAGACTCTGCTC 473

RESULT 185
US-09-949-016-204542
; Sequence 204542, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 204542
; LENGTH: 601
```

```

; TYPE: DNA
; ORGANISM: Human
US-09-949-016-204542

Query Match
Best Local Similarity 100.0%; Score 44; DB 3; Length 601;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TCCCACTGCACTCCAGCTCGGCAACAGCAAGACTCTGCTC 3122
DB 430 TCCCACTGCACTCCAGCTCGGCAACAGCAAGACTCTGCTC 473

RESULT 186
US-09-949-002-10588/C
; Sequence 10588, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10588
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-10588

Query Match
Best Local Similarity 100.0%; Score 44; DB 3; Length 601;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TCCCACTGCACTCCAGCTCGGCAACAGCAAGACTCTGCTC 3122
DB 168 TCCCACTGCACTCCAGCTCGGCAACAGCAAGACTCTGCTC 125

RESULT 187
US-09-641-638-56
; Sequence 56, Application US/09641638
; Patent No. 6432648
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Ilya
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIALLYLIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; FILE REFERENCE: GENSET.051CP1
; CURRENT APPLICATION NUMBER: US/09/641,638
; CURRENT FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 56
; LENGTH: 956
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURES:
```

```
; NAME/KEY: allele
; LOCATION: 478
; OTHER INFORMATION: 12-220-48 : polymorphic base G or A
; NAME/KEY: misc_binding
; LOCATION: 458..477
; OTHER INFORMATION: 12-220-48.mis1, potential
; NAME/KEY: misc_binding
; LOCATION: 479..498
; OTHER INFORMATION: 12-220-48.mis2, potential complement
; NAME/KEY: primer_bind
; LOCATION: 505..525
; OTHER INFORMATION: upstream amplification primer, complement
; NAME/KEY: primer_bind
; LOCATION: 76..96
; OTHER INFORMATION: downstream amplification primer
; NAME/KEY: misc_binding
; LOCATION: 466..490
; OTHER INFORMATION: 12-220-48 potential probe
US-09-641-638-56

Query Match      1.4%; Score 44; DB 3; Length 956;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3073 AGATTGTGCACTGCACTCCAGCTGGGCAAGAGAGACTC 3116
Db      802 AGATTGTGCACTGCACTCCAGCTGGGCAAGAGAGACTC 845

RESULT 188
US-10-170-097-56
; Sequence 56: Application US/10170097
; Patent No. 6794143
; GENERAL INFORMATION:
; APPLICANT: Blumenfeld, Marta
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Chumakov, Ilya
; APPLICANT: Cohen, Annick
; TITLE OF INVENTION: BIOMETRIC MARKERS DERIVED FROM GENOMIC REGIONS CARRYING
; TITLE OF INVENTION: GENES INVOLVED IN ARACHIDONIC ACID METABOLISM
; FILE REFERENCE: GEN-T114XC2D1
; CURRENT FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/641,638
; PRIOR FILING DATE: 2000-08-16
; PRIOR APPLICATION NUMBER: US 09/502,330
; PRIOR FILING DATE: 2000-02-11
; PRIOR APPLICATION NUMBER: US 60/133,200
; PRIOR FILING DATE: 1999-05-07
; PRIOR APPLICATION NUMBER: US 09/275,267
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: US 60/119,917
; PRIOR FILING DATE: 1999-02-12
; NUMBER OF SEQ ID NOS: 1304
; SOFTWARE: Patent.pm
; SEQ ID NO 56
; LENGTH: 956
; TYPE: DNA
; ORGANISM: Homo Sapiens
; FEATURE:
; NAME/KEY: allele
; LOCATION: 478
; OTHER INFORMATION: 12-220-48 : polymorphic base G or A
; FEATURE:
; FILE REFERENCE:
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; NAME/KEY: primer_bind
```

```
; LOCATION: 505..525
; OTHER INFORMATION: upstream amplification primer, complement
; FEATURE:
; NAME/KEY: primer_bind
; LOCATION: 76..96
; OTHER INFORMATION: downstream amplification primer
; FEATURE:
; NAME/KEY: misc_binding
; LOCATION: 466..490
; OTHER INFORMATION: 12-220-48 potential probe
US-10-170-097-56

Query Match      1.4%; Score 44; DB 3; Length 956;
Best Local Similarity 100.0%; Pred. No. 1.9e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3073 AGATTGTGCACTGCACTCCAGCTGGGCAAGAGAGACTC 3116
Db      802 AGATTGTGCACTGCACTCCAGCTGGGCAAGAGAGACTC 845

RESULT 189
US-09-949-016-11999
; Sequence 11999: Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 11999
; LENGTH: 7480
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-11999

Query Match      1.4%; Score 44; DB 3; Length 7480;
Best Local Similarity 100.0%; Pred. No. 1.8e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3079 TGGCACTGCACTCCAGCTGGGCAAGAGAGACTCTGCTC 3122
Db      247 TGGCACTGCACTCCAGCTGGGCAAGAGAGACTCTGCTC 290

RESULT 190
US-09-949-016-16486
; Sequence 16486: Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
```

SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 16486  
LENGTH: 7567  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-16486

Query Match 1.4%; Score 44; DB 3; Length 7567;  
Best Local Similarity 100.0%; Pred. No. 1.8e-10;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3079 TGCACCTGCACTCCAGCTCGGGCAACAGCAAGACTCTGTCTC 3122  
DB 334 TGCACCTGCACTCCAGCTCGGGCAACAGCAAGACTCTGTCTC 377

RESULT 191  
US-09-318-448-5/c  
Sequence 5, Application US/09318448  
Patent No. 6210950  
GENERAL INFORMATION:  
APPLICANT: Johnson, William G.  
APPLICANT: Stencos, Edward S.  
TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING  
FILE REFERENCE: 601-1-057  
CURRENT APPLICATION NUMBER: US/09/318,448  
CURRENT FILING DATE: 1999-05-25  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 5  
LENGTH: 7720  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-318-448-5

Query Match 1.4%; Score 44; DB 3; Length 7720;  
Best Local Similarity 100.0%; Pred. No. 1.8e-10;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACTGCACTCCAGCTCGGGCAACAGCAAGACTC 3116  
DB 5249 AGATTGTGCACTGCACTCCAGCTCGGGCAACAGCAAGACTC 5206

RESULT 192  
US-09-577-266-5/c  
Sequence 5, Application US/09577266  
Patent No. 6912492  
GENERAL INFORMATION:  
APPLICANT: Johnson, William G.  
APPLICANT: Stencos, Edward S.  
TITLE OF INVENTION: METHODS FOR DIAGNOSING, PREVENTING, AND TREATING  
FILE REFERENCE: 601-1-057N  
CURRENT APPLICATION NUMBER: US/09/577,266  
CURRENT FILING DATE: 2000-05-23  
PRIOR APPLICATION NUMBER: 60/136,198  
PRIOR FILING DATE: 1999-05-25  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 5  
LENGTH: 7720  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-577-266-5

Query Match 1.4%; Score 44; DB 3; Length 7720;  
Best Local Similarity 100.0%; Pred. No. 1.8e-10;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACTGCACTCCAGCTCGGGCAACAGCAAGACTC 3116  
|||||

DB 5249 AGATTGTGCACTGCACTCCAGCTCGGGCAACAGCAAGACTC 5206

RESULT 193  
US-09-949-016-17443/c  
Sequence 17443, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 17443  
LENGTH: 10720  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-17443

Query Match 1.4%; Score 44; DB 3; Length 10720;  
Best Local Similarity 100.0%; Pred. No. 1.7e-10;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACTGCACTCCAGCTCGGGCAACAGCAAGACTC 3116  
DB 3805 AGATTGTGCACTGCACTCCAGCTCGGGCAACAGCAAGACTC 3762

RESULT 194  
US-09-949-016-12354  
Sequence 12354, Application US/09949016  
Patent No. 6812339  
GENERAL INFORMATION:  
APPLICANT: VENTER, J. Craig et al.  
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
FILE REFERENCE: CL001307  
CURRENT APPLICATION NUMBER: US/09/949,016  
CURRENT FILING DATE: 2000-04-14  
PRIOR APPLICATION NUMBER: 60/241,755  
PRIOR FILING DATE: 2000-10-20  
PRIOR APPLICATION NUMBER: 60/237,768  
PRIOR FILING DATE: 2000-10-03  
PRIOR APPLICATION NUMBER: 60/231,498  
PRIOR FILING DATE: 2000-09-08  
NUMBER OF SEQ ID NOS: 207012  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 12354  
LENGTH: 15116  
TYPE: DNA  
ORGANISM: Human  
US-09-949-016-12354

Query Match 1.4%; Score 44; DB 3; Length 15116;  
Best Local Similarity 100.0%; Pred. No. 1.7e-10;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2899 GATCACTGAGGCGAGATTGAGACCGAGCTGGGCAACATAG 2942  
DB 8745 GATCACTGAGGCGAGATTGAGACCGAGCTGGGCAACATAG 8788

RESULT 195  
US-09-949-016-16260

```
; Sequence 16260, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 16260
; LENGTH: 15117
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16260
```

```
Query Match 1.4%; Score 44; DB 3; Length 15117;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2899 GATCAGCTGAGGCGGAGATTGTGAGACCGCTGGCCCAACATG 2942
DB 8745 GATCAGCTGAGGCGGAGATTGTGAGACCGCTGGCCCAACATG 8788
```

```
RESULT 196
US-09-949-002-592
; Sequence 592, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; FILE REFERENCE: CL000790
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 592
; LENGTH: 18651
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-592
```

```
Query Match 1.4%; Score 44; DB 3; Length 18651;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3073 AGATTGTGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
DB 14614 AGATTGTGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 14657
```

```
RESULT 197
US-09-949-002-786
; Sequence 786, Application US/09949002
; Patent No. 6900016
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION
; FILE REFERENCE: CL000790
```

```
; CURRENT APPLICATION NUMBER: US/09/949,002
; CURRENT FILING DATE: 2000-01-28
; PRIOR APPLICATION NUMBER: 60/231,401
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 10823
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 786
; LENGTH: 18682
; TYPE: DNA
; ORGANISM: Human
US-09-949-002-786
```

```
Query Match 1.4%; Score 44; DB 3; Length 18682;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3073 AGATTGTGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
DB 14634 AGATTGTGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 14677
```

```
RESULT 198
US-09-949-016-16523/c
; Sequence 16523, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 16523
; LENGTH: 18891
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(18891)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-16523
```

```
Query Match 1.4%; Score 44; DB 3; Length 18891;
Best Local Similarity 100.0%; Pred. No. 1.7e-10;
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3073 AGATTGTGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
DB 14166 AGATTGTGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 14123
```

```
RESULT 199
US-09-949-016-15113
; Sequence 15113, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH INFLAMMATORY AUTOIMMUNE DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
```

;; PRIOR FILING DATE: 2000-10-03  
;; PRIOR APPLICATION NUMBER: 60/231,498  
;; PRIOR FILING DATE: 2000-09-08  
;; NUMBER OF SEQ ID NOS: 207012  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 15113  
;; LENGTH: 19389  
;; TYPE: DNA  
;; ORGANISM: Human  
US-09-949-016-15113

Query Match 1.4%; Score 44; DB 3; Length 19389;  
Best Local Similarity 100.0%; Pred. No. 1.7e-10;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTC 3116  
DB 14852 AGATTGTGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTC 14895

## RESULT 200

US-09-949-016-15267  
; Sequence 15267, Application US/09949016  
; Patent No. 6812339  
; GENERAL INFORMATION:  
; APPLICANT: VENTER, J. Craig et al.  
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED  
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001307  
; CURRENT APPLICATION NUMBER: US/09/949,016  
; CURRENT FILING DATE: 2000-04-14  
; PRIOR APPLICATION NUMBER: 60/241,755  
; PRIOR FILING DATE: 2000-10-20  
; PRIOR APPLICATION NUMBER: 60/237,768  
; PRIOR FILING DATE: 2000-10-03  
; PRIOR APPLICATION NUMBER: 60/231,498  
-; PRIOR FILING DATE: 2000-09-08  
; NUMBER OF SEQ ID NOS: 207012  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 15267  
; LENGTH: 19974  
; TYPE: DNA  
; ORGANISM: Human  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(19974)  
; OTHER INFORMATION: n = A,T,C or G  
US-09-949-016-15267

Query Match 1.4%; Score 44; DB 3; Length 19974;  
Best Local Similarity 100.0%; Pred. No. 1.7e-10;  
Matches 44; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTC 3116  
DB 15162 AGATTGTGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTC 15205

Search completed: May 11, 2006, 06:00:44  
Job time : 560 secs





GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2006, 05:51:35 ; Search time 2382 Seconds  
(without alignments)  
10838.370 Million cell updates/sec

Title: US-10-760-320A-102

Perfect score: 3122

Sequence: 1 actagaggtcggggttagcgcc.....acagagcaagactctgtctc 3122

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 9793542 seqs, 413468905 residues

Word size : 1

Total number of hits satisfying chosen parameters: 19584538

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 500 summaries

Database : Published Applications NA Main:\*

1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*  
3: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*  
4: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
6: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
8: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	748	24.0	850	9	US-10-450-763-29532
2	657	21.0	708	6	US-10-029-386-20699
3	524	16.8	524	6	US-10-029-386-6988
4	509	16.3	560	4	US-09-925-065A-740956
5	506	16.2	557	4	US-09-925-065A-769508
6	505	16.2	556	4	US-09-925-065A-736351
7	504	16.1	555	4	US-09-925-065A-737120
8	241	7.9	1349	9	US-10-450-763-8312
9	60	1.9	60	3	US-09-908-975-23330
10	53	1.7	175737	8	US-10-723-860-1753
11	53	1.7	175737	9	US-10-723-860-1753
12	52	1.7	428	8	US-10-674-1244-76
13	52	1.7	559	4	US-09-925-065A-531947
14	52	1.7	568	4	US-10-027-632-41688
15	52	1.7	574	4	US-09-925-065A-768118
16	52	1.7	592	4	US-09-925-065A-59682
17	52	1.7	617	4	US-09-925-065A-386361
18	52	1.7	631	5	US-10-027-632-61987
19	52	1.7	631	5	US-10-027-632-61987
20	52	1.7	631	5	US-10-027-632-61989
21	52	1.7	631	5	US-10-027-632-61989
22	52	1.7	631	6	US-10-027-632-61987
23	52	1.7	631	6	US-10-027-632-61988

24	52	1.7	631	6	US-10-027-632-61989	Sequence 61989, A
25	52	1.7	695	4	US-09-925-065A-708701	Sequence 708701, A
26	52	1.7	1437	9	US-10-450-763-14141	Sequence 14141, A
27	52	1.7	9620	3	US-09-764-881-8895	Sequence 8895, A
28	52	1.7	92139	6	US-09-918-686-1	Sequence 1, A
29	52	1.7	92139	6	US-10-353-150-1	Sequence 1, A
30	52	1.7	130320	6	US-10-408-168-1	Sequence 1, A
31	52	1.7	136726	6	US-10-085-117-244	Sequence 244, A
32	52	1.7	150437	6	US-10-981-277-50	Sequence 12, A
33	52	1.7	221000	6	US-10-174-014-12	Sequence 12, A
34	52	1.7	233380	5	US-10-087-192-652	Sequence 652, A
35	52	1.7	398287	7	US-10-741-601-5719	Sequence 5719, A
36	52	1.7	398287	8	US-10-741-600-17839	Sequence 17839, A
37	51	1.6	385	4	US-09-925-065A-111249	Sequence 111249, A
38	51	1.6	561	4	US-09-925-065A-414933	Sequence 414933, A
39	51	1.6	622	5	US-10-027-632-54553	Sequence 54553, A
40	51	1.6	622	5	US-10-027-632-55034	Sequence 55034, A
41	51	1.6	622	5	US-10-027-632-55282	Sequence 55282, A
42	51	1.6	622	5	US-10-027-632-55384	Sequence 55384, A
43	51	1.6	622	5	US-10-027-632-180536	Sequence 180536, A
44	51	1.6	622	5	US-10-027-632-314485	Sequence 314485, A
45	51	1.6	622	6	US-10-027-632-54553	Sequence 54553, A
46	51	1.6	622	6	US-10-027-632-55034	Sequence 55034, A
47	51	1.6	622	6	US-10-027-632-55282	Sequence 55282, A
48	51	1.6	622	6	US-10-027-632-55384	Sequence 55384, A
49	51	1.6	622	6	US-10-027-632-180536	Sequence 180536, A
50	51	1.6	622	6	US-10-027-632-314485	Sequence 314485, A
51	51	1.6	832	3	US-09-925-298-220	Sequence 220, A
52	51	1.6	832	3	US-10-102-806-220	Sequence 220, A
53	51	1.6	2791	9	US-10-450-763-29534	Sequence 29534, A
54	51	1.6	23139	9	US-10-737-082-22	Sequence 22, A
55	51	1.6	23139	9	US-10-765-790-22	Sequence 22, A
56	51	1.6	68200	9	US-10-840-159-3	Sequence 3, A
57	51	1.6	215221	5	US-10-087-192-1360	Sequence 1360, A
58	50	1.6	288	3	US-09-973-278-870	Sequence 870, A
59	50	1.6	432	9	US-10-450-763-29529	Sequence 29529, A
60	50	1.6	571	4	US-09-925-065A-84883	Sequence 84883, A
61	50	1.6	572	4	US-09-925-065A-785967	Sequence 785967, A
62	50	1.6	625	4	US-09-925-065A-894777	Sequence 894777, A
63	50	1.6	630	4	US-09-925-065A-892664	Sequence 892664, A
64	50	1.6	633	4	US-09-925-065A-892405	Sequence 892405, A
65	50	1.6	633	4	US-09-925-065A-925214	Sequence 925214, A
66	50	1.6	633	4	US-09-925-065A-949277	Sequence 949277, A
67	50	1.6	676	4	US-09-925-065A-917225	Sequence 917225, A
68	50	1.6	676	4	US-09-925-065A-944569	Sequence 944569, A
69	50	1.6	714	5	US-10-027-632-158850	Sequence 158850, A
70	50	1.6	714	5	US-10-027-632-158850	Sequence 158850, A
71	50	1.6	714	6	US-10-027-632-158850	Sequence 158850, A
72	50	1.6	739	4	US-09-925-065A-934608	Sequence 934608, A
73	50	1.6	742	4	US-09-925-065A-941367	Sequence 941367, A
74	50	1.6	1986	5	US-10-027-632-265269	Sequence 265269, A
75	50	1.6	1986	6	US-10-027-632-265269	Sequence 265269, A
76	50	1.6	8705	7	US-10-374-979-10	Sequence 10, A
77	50	1.6	8705	7	US-10-182-938A-10	Sequence 10, A
78	50	1.6	8705	8	US-10-731-733-10	Sequence 10, A
79	50	1.6	8705	8	US-10-477-238A-10	Sequence 10, A
80	50	1.6	8705	8	US-10-680-287A-10	Sequence 10, A
81	50	1.6	8705	9	US-10-477-177-10	Sequence 10, A
82	50	1.6	8705	9	US-10-834-337-10	Sequence 10, A
83	50	1.6	31474	3	US-09-764-878-3149	Sequence 3149, A
84	50	1.6	31474	3	US-09-764-878-3149	Sequence 3149, A
85	50	1.6	32189	5	US-10-079-854-379	Sequence 379, A
86	50	1.6	32189	5	US-10-079-854-379	Sequence 379, A
87	50	1.6	32193	3	US-09-764-877-2623	Sequence 2623, A
88	50	1.6	32193	3	US-10-216-464-37	Sequence 37, A
89	50	1.6	32221	6	US-10-242-515-2623	Sequence 2623, A
90	50	1.6	32221	5	US-09-764-878-377	Sequence 377, A
91	50	1.6	33226	5	US-10-079-854-377	Sequence 377, A
92	50	1.6	33226	5	US-10-085-117-190	Sequence 190, A
93	50	1.6	56510	8	US-10-741-600-17637	Sequence 17637, A
94	50	1.6	56510	8	US-10-087-192-574	Sequence 574, A
95	50	1.6	156643	5	US-10-087-192-574	Sequence 1408, A
96	50	1.6	202814	8	US-10-719-993-6812	Sequence 6812, A
97	7	1.6	276820	7	US-10-271-416-9	Sequence 9, A

C 97	49	1.6	420	4	US-09-925-065A-140670	Sequence 140670,	170	46	1.5	689	4	US-09-925-065A-852555	Sequence 852555,
C 98	49	1.6	532	4	US-09-925-065A-140671	Sequence 140671,	171	46	1.5	689	4	US-09-925-065A-852556	Sequence 852556,
C 99	49	1.6	552	4	US-09-925-065A-186091	Sequence 186091, A	172	46	1.5	756	5	US-10-027-632-144717	Sequence 144717,
C 100	49	1.6	552	4	US-09-925-065A-18610	Sequence 18610, A	173	46	1.5	756	5	US-10-027-632-144718	Sequence 144718,
C 101	49	1.6	552	4	US-09-925-065A-18611	Sequence 18611, A	174	46	1.5	756	5	US-10-027-632-144719	Sequence 144719,
C 102	49	1.6	580	5	US-10-027-632-80386	Sequence 80386, A	175	46	1.5	756	6	US-10-027-632-144717	Sequence 144717,
C 103	49	1.6	580	5	US-10-027-632-80387	Sequence 80387, A	176	46	1.5	756	6	US-10-027-632-144718	Sequence 144718,
C 104	49	1.6	580	5	US-10-027-632-80388	Sequence 80388, A	177	46	1.5	756	6	US-10-027-632-144719	Sequence 144719,
C 105	49	1.6	580	5	US-10-027-632-80389	Sequence 80389, A	178	46	1.5	840	4	US-09-925-065A-938958	Sequence 938958,
C 106	49	1.6	695	4	US-09-925-065A-708700	Sequence 708700, A	179	46	1.5	840	4	US-09-925-065A-938959	Sequence 938959,
C 107	49	1.6	750	4	US-09-925-065A-932035	Sequence 932035,	180	46	1.5	840	4	US-09-925-065A-938960	Sequence 938960,
C 108	49	1.6	750	4	US-09-925-065A-932036	Sequence 932036,	181	46	1.5	1105	4	US-09-925-065A-711495	Sequence 711495,
C 109	49	1.6	3170	5	US-10-027-632-113573	Sequence 113573,	182	46	1.5	1364	4	US-09-925-065A-678054	Sequence 678054,
C 110	49	1.6	3170	5	US-10-027-632-113574	Sequence 113574,	183	46	1.5	1635	4	US-09-925-065A-689300	Sequence 689300,
C 111	49	1.6	3170	5	US-10-027-632-113575	Sequence 113575,	184	46	1.5	2227	4	US-09-925-065A-677066	Sequence 677066,
C 112	49	1.6	3170	5	US-10-027-632-113576	Sequence 113576,	185	46	1.5	2798	5	US-10-027-632-260094	Sequence 260094,
C 113	48	1.5	460	5	US-10-027-632-182264	Sequence 182264,	186	46	1.5	2798	6	US-10-027-632-260094	Sequence 260094,
C 114	48	1.5	460	5	US-10-027-632-319423	Sequence 319423,	187	46	1.5	7001	6	US-10-027-632-260094	Sequence 260094,
C 115	48	1.5	460	5	US-10-027-632-319423	Sequence 319423,	188	46	1.5	7001	6	US-10-027-632-260094	Sequence 260094,
C 116	48	1.5	460	5	US-10-027-632-319423	Sequence 319423,	189	46	1.5	13409	3	US-09-764-891-9601	Sequence 9601, Ap
C 117	48	1.5	716	5	US-10-027-632-113626	Sequence 113626,	190	46	1.5	13409	5	US-10-205-428-930	Sequence 930, Ap
C 118	48	1.5	716	5	US-10-027-632-113626	Sequence 113626,	191	46	1.5	15601	3	US-09-764-847-1916	Sequence 1916, Ap
C 119	48	1.5	52242	7	US-10-052-482-1172	Sequence 172, App	192	46	1.5	15601	3	US-10-092-154-1916	Sequence 1916, Ap
C 120	48	1.5	53779	9	US-10-737-082-85	Sequence 85, App1	193	46	1.5	25001	6	US-10-187-6594-11	Sequence 11, App1
C 121	48	1.5	53779	9	US-10-765-790-85	Sequence 85, App1	194	46	1.5	25001	6	US-10-187-6594-11	Sequence 11, App1
C 122	48	1.5	141121	8	US-10-741-600-17651	Sequence 17651, A	195	46	1.5	57095	8	US-10-719-993-7030	Sequence 7030, Ap
C 123	48	1.5	181684	5	US-10-087-192-790	Sequence 790, App	196	46	1.5	57095	8	US-10-322-281-526	Sequence 526, App
C 124	47	1.5	440	8	US-10-357-930-16322	Sequence 16322, A	197	46	1.5	55914	7	US-10-741-601-5519	Sequence 5519, Ap
C 125	47	1.5	516	8	US-10-357-930-46148	Sequence 46148, A	198	46	1.5	95011	8	US-10-719-993-6671	Sequence 6671, Ap
C 126	47	1.5	516	8	US-09-925-065A-440136	Sequence 440136,	199	46	1.5	95011	8	US-10-719-993-6671	Sequence 6671, Ap
C 127	47	1.5	534	4	US-09-925-065A-440137	Sequence 440137,	200	46	1.5	95011	8	US-10-052-482-222	Sequence 222
C 128	47	1.5	556	4	US-09-925-065A-623360	Sequence 623360,	201	46	1.5	99250	9	US-10-840-530-4	Sequence 4, App1
C 129	47	1.5	3231	5	US-10-027-632-115306	Sequence 115306,	202	46	1.5	101782	7	US-10-741-601-5661	Sequence 5661, Ap
C 130	47	1.5	3231	5	US-10-027-632-115307	Sequence 115307,	203	46	1.5	103391	9	US-10-981-227-54	Sequence 22, App1
C 131	47	1.5	3231	5	US-10-027-632-115308	Sequence 115308,	204	46	1.5	141463	5	US-10-087-192-22	Sequence 22, App1
C 132	47	1.5	3231	6	US-10-027-632-115306	Sequence 115306,	205	46	1.5	168821	7	US-10-087-192-622	Sequence 622, App
C 133	47	1.5	3231	6	US-10-027-632-115307	Sequence 115307,	206	46	1.5	181259	5	US-10-456-930-2	Sequence 2, App1
C 134	47	1.5	3231	6	US-10-027-632-115308	Sequence 115308,	207	46	1.5	181259	5	US-10-085-117-142	Sequence 142, App
C 135	46	1.5	201	8	US-10-719-993-51173	Sequence 51173, A	208	46	1.5	256157	6	US-10-087-192-1204	Sequence 1204, App
C 136	46	1.5	322	8	US-10-674-124A-25317	Sequence 25317, A	209	46	1.5	256157	7	US-10-322-281-776	Sequence 776, App
C 137	46	1.5	380	8	US-10-914-037-282	Sequence 282, App	210	46	1.5	276276	5	US-10-087-192-754	Sequence 754, App
C 138	46	1.5	405	3	US-09-867-701-6376	Sequence 6376, App	211	46	1.5	347814	7	US-10-322-666-76	Sequence 76, App1
C 139	46	1.5	440	8	US-10-674-124A-24232	Sequence 24232, A	212	46	1.5	410846	9	US-10-481-613-1	Sequence 1, App1
C 140	46	1.5	458	3	US-09-814-353-17260	Sequence 17260, A	213	46	1.5	71651	8	US-10-719-993-6617	Sequence 6617, A
C 141	46	1.5	482	8	US-10-674-124A-16717	Sequence 16717, A	214	46	1.5	744802	6	US-10-292-798-1369	Sequence 1369, Ap
C 142	46	1.5	497	8	US-10-357-930-60554	Sequence 60554, A	215	46	1.5	1445	3	US-09-764-868-1298	Sequence 1298, Ap
C 143	46	1.5	535	5	US-10-027-632-91304	Sequence 91304, A	216	45	1.4	1445	3	US-09-764-875-1230	Sequence 1230, Ap
C 144	46	1.5	535	5	US-10-027-632-91305	Sequence 91305, A	217	45	1.4	425	5	US-10-027-632-280871	Sequence 280871,
C 145	46	1.5	535	5	US-10-027-632-91306	Sequence 91306, A	218	45	1.4	425	5	US-10-027-632-280871	Sequence 280871,
C 146	46	1.5	535	6	US-10-027-632-91304	Sequence 91304, A	219	45	1.4	492	3	US-09-918-995-28327	Sequence 28327, A
C 147	46	1.5	535	6	US-10-027-632-91305	Sequence 91305, A	220	45	1.4	541	5	US-10-027-632-193287	Sequence 193287,
C 148	46	1.5	535	6	US-10-027-632-91306	Sequence 91306, A	221	45	1.4	541	5	US-10-027-632-193288	Sequence 193288,
C 149	46	1.5	535	4	US-09-925-065A-124429	Sequence 124429,	222	45	1.4	541	6	US-10-027-632-193288	Sequence 193288,
C 150	46	1.5	606	4	US-09-925-065A-843811	Sequence 843811,	223	45	1.4	541	6	US-10-027-632-193288	Sequence 193288,
C 151	46	1.5	610	4	US-09-925-065A-801712	Sequence 801712,	224	45	1.4	545	4	US-09-925-065A-765896	Sequence 765896,
C 152	46	1.5	610	4	US-09-925-065A-48663	Sequence 48663, A	225	45	1.4	545	4	US-09-925-065A-799095	Sequence 799095,
C 153	46	1.5	610	4	US-09-925-065A-65044	Sequence 65044, A	226	45	1.4	549	4	US-09-925-065A-615308	Sequence 615308,
C 154	46	1.5	614	4	US-09-925-065A-47410	Sequence 47410, A	227	45	1.4	552	5	US-10-027-632-68940	Sequence 68940, A
C 155	46	1.5	614	4	US-09-925-065A-47411	Sequence 47411, A	228	45	1.4	552	5	US-10-027-632-284750	Sequence 284750, A
C 156	46	1.5	615	4	US-09-925-065A-930543	Sequence 930543,	229	45	1.4	552	6	US-10-027-632-68940	Sequence 68940, A
C 157	46	1.5	615	4	US-09-925-065A-930544	Sequence 930544,	230	45	1.4	552	6	US-10-027-632-284750	Sequence 284750,
C 158	46	1.5	616	4	US-09-925-065A-870312	Sequence 870312,	231	45	1.4	558	5	US-10-027-632-288369	Sequence 288369,
C 159	46	1.5	621	5	US-10-027-632-115128	Sequence 115128,	232	45	1.4	558	6	US-10-027-632-288369	Sequence 288369,
C 160	46	1.5	621	5	US-10-027-632-115129	Sequence 115129,	233	45	1.4	559	4	US-09-925-065A-931948	Sequence 931948,
C 161	46	1.5	621	6	US-10-027-632-115128	Sequence 115128,	234	45	1.4	563	4	US-09-925-065A-12438	Sequence 12438, A
C 162	46	1.5	621	6	US-10-027-632-115129	Sequence 115129,	235	45	1.4	568	4	US-09-925-065A-31088	Sequence 31088, A
C 163	46	1.5	654	4	US-09-925-065A-554324	Sequence 554324,	236	45	1.4	568	4	US-09-925-065A-318452	Sequence 318452,
C 164	46	1.5	654	4	US-09-925-065A-759345	Sequence 759345,	237	45	1.4	590	4	US-09-925-065A-927765	Sequence 927765,
C 165	46	1.5	659	4	US-09-925-065A-790726	Sequence 790726,	238	45	1.4	592	4	US-09-925-065A-927765	Sequence 927765,
C 166	46	1.5	659	4	US-09-925-065A-790727	Sequence 790727,	239	45	1.4	592	4	US-09-925-065A-927765	Sequence 927765,
C 167	46	1.5	661	4	US-09-925-065A-771454	Sequence 771454,	240	45	1.4	592	4	US-09-925-065A-949474	Sequence 949474,
C 168	46	1.5	672	4	US-09-925-065A-737520	Sequence 737520,	241	45	1.4	596	4	US-09-925-065A-695847	Sequence 695847,
C 169	46	1.5	680	4	US-09-925-065A-737227	Sequence 737227,	242	45	1.4	598	4	US-09-925-065A-10892	Sequence 10892, A

243	45	1.4	601	9	US-10-893-315-1202	Sequence 1202, Ap	316	45	1.4	115755	8	US-10-719-993-6946	Sequence 6946, Ap
C 244	45	1.4	604	4	US-09-925-065A-160227	Sequence 160227,	317	45	1.4	115661	8	US-10-723-860-B-49	Sequence 409, App
C 245	45	1.4	604	4	US-09-925-065A-799096	Sequence 799096,	C 318	45	1.4	149671	6	US-10-236-031B-53	Sequence 53, App
C 246	45	1.4	605	4	US-09-925-065A-795229	Sequence 795229,	C 319	45	1.4	150573	9	US-10-981-227-56	Sequence 56, Appl
C 247	45	1.4	605	4	US-09-925-065A-850785	Sequence 850785,	C 320	45	1.4	191584	7	US-10-322-281-800	Sequence 800, Appl
C 248	45	1.4	612	5	US-10-927-632-295528	Sequence 295528,	C 321	45	1.4	300000	6	US-10-262-552-33	Sequence 33, Appl
C 249	45	1.4	612	6	US-10-927-632-295528	Sequence 295528,	C 322	45	1.4	300000	7	US-10-703-210-33	Sequence 33, Appl
C 250	45	1.4	624	4	US-09-925-065A-864941	Sequence 864941,	C 323	45	1.4	366803	8	US-10-719-993-6805	Sequence 6805, Ap
C 251	45	1.4	627	4	US-09-925-065A-839499	Sequence 839499,	C 324	45	1.4	368604	8	US-09-949-654-3	Sequence 3, Appl
C 252	45	1.4	629	5	US-10-927-632-14053	Sequence 14053, A	C 325	45	1.4	567864	8	US-10-699-156-3	Sequence 3, Appl
C 253	45	1.4	629	6	US-10-927-632-14053	Sequence 14053, A	C 326	45	1.4	786452	8	US-10-719-993-6822	Sequence 6822, Ap
C 254	45	1.4	642	5	US-10-927-632-105085	Sequence 105085,	C 327	45	1.4	1691139	5	US-10-067-514-1	Sequence 1, Appl
C 255	45	1.4	642	5	US-10-927-632-135499	Sequence 135499,	C 328	45	1.4	1691139	7	US-10-419-723-1	Sequence 1, Appl
C 256	45	1.4	642	5	US-10-927-632-135499	Sequence 135499,	C 329	45	1.4	1691139	7	US-10-255-120-1	Sequence 1, Appl
C 257	45	1.4	642	5	US-10-927-632-135499	Sequence 135499,	C 330	45	1.4	1691140	9	US-10-868-397-1	Sequence 1, Appl
C 258	45	1.4	642	6	US-10-927-632-105085	Sequence 105085,	C 331	44	1.4	98	9	US-10-708-204-6436	Sequence 6436, Ap
C 259	45	1.4	642	6	US-10-927-632-135499	Sequence 135499,	C 332	44	1.4	116	3	US-09-984-425-588	Sequence 588, App
C 260	45	1.4	642	6	US-10-927-632-135499	Sequence 135499,	C 333	44	1.4	129	7	US-10-242-535A-52981	Sequence 52981, A
C 261	45	1.4	642	6	US-10-927-632-135500	Sequence 135500,	C 334	44	1.4	129	7	US-10-085-783A-52981	Sequence 52981, A
C 262	45	1.4	652	5	US-10-927-632-184845	Sequence 184845,	C 335	44	1.4	182	3	US-09-764-887-597	Sequence 597, App
C 263	45	1.4	652	5	US-10-927-632-184846	Sequence 184846,	C 336	44	1.4	182	5	US-10-073-961-597	Sequence 597, App
C 264	45	1.4	652	6	US-10-927-632-184845	Sequence 184845,	C 337	44	1.4	201	8	US-10-719-993-12850	Sequence 12850, A
C 265	45	1.4	652	6	US-10-927-632-184846	Sequence 184846,	C 338	44	1.4	250	5	US-10-027-633-22856	Sequence 22856, A
C 266	45	1.4	660	5	US-10-927-632-216907	Sequence 216907,	C 339	44	1.4	250	6	US-10-027-633-22856	Sequence 22856, A
C 267	45	1.4	660	5	US-10-927-632-216908	Sequence 216908,	C 340	44	1.4	293	8	US-10-674-124A-23859	Sequence 23859, A
C 268	45	1.4	660	6	US-10-927-632-216907	Sequence 216907,	C 341	44	1.4	312	8	US-10-674-124A-932	Sequence 932, App
C 269	45	1.4	660	6	US-10-927-632-216908	Sequence 216908,	C 342	44	1.4	423	4	US-09-925-065A-566681	Sequence 566681,
C 270	45	1.4	662	4	US-09-925-065A-675355	Sequence 675355,	C 343	44	1.4	423	4	US-09-925-065A-566682	Sequence 566682,
C 271	45	1.4	662	4	US-09-925-065A-675356	Sequence 675356,	C 344	44	1.4	423	5	US-10-027-633-36530	Sequence 36530, A
C 272	45	1.4	677	5	US-10-927-632-121998	Sequence 121998,	C 345	44	1.4	423	5	US-10-027-633-36531	Sequence 36531, A
C 273	45	1.4	677	6	US-10-927-632-121998	Sequence 121998,	C 346	44	1.4	423	5	US-10-027-633-80045	Sequence 80045, A
C 274	45	1.4	702	5	US-10-927-632-145794	Sequence 145794,	C 347	44	1.4	423	5	US-10-027-633-80046	Sequence 80046, A
C 275	45	1.4	702	6	US-10-927-632-145794	Sequence 145794,	C 348	44	1.4	423	5	US-10-027-633-109592	Sequence 109592,
C 276	45	1.4	705	4	US-10-927-632-149399	Sequence 149399,	C 349	44	1.4	423	6	US-10-027-633-109592	Sequence 109592,
C 277	45	1.4	705	6	US-10-927-632-149399	Sequence 149399,	C 350	44	1.4	423	6	US-10-027-633-36530	Sequence 36530, A
C 278	45	1.4	793	4	US-09-925-065A-13524	Sequence 13524, A	C 351	44	1.4	423	6	US-10-027-633-36531	Sequence 36531, A
C 279	45	1.4	793	4	US-09-925-065A-77588	Sequence 77588, A	C 352	44	1.4	426	3	US-10-843-645A-5054	Sequence 80045, A
C 280	45	1.4	794	5	US-10-925-065A-88887	Sequence 88887, A	C 353	44	1.4	426	3	US-10-843-645A-5054	Sequence 80046, A
C 281	45	1.4	896	6	US-10-925-065A-88887	Sequence 88887, A	C 354	44	1.4	423	6	US-10-027-633-109591	Sequence 109591,
C 282	45	1.4	896	6	US-10-925-065A-88887	Sequence 88887, A	C 355	44	1.4	423	6	US-10-027-633-109592	Sequence 109592,
C 283	45	1.4	1027	4	US-09-925-065A-85535	Sequence 85535, A	C 356	44	1.4	426	3	US-10-954-455-2027	Sequence 2027, Ap
C 284	45	1.4	1027	4	US-09-925-065A-85536	Sequence 85536, A	C 357	44	1.4	426	3	US-10-954-455-2027	Sequence 2027, Ap
C 285	45	1.4	1302	4	US-09-925-065A-553464	Sequence 553464,	C 358	44	1.4	464	4	US-09-925-065A-602339	Sequence 602339,
C 286	45	1.4	1302	4	US-09-925-065A-553464	Sequence 553464,	C 359	44	1.4	470	4	US-09-925-065A-602339	Sequence 602339,
C 287	45	1.4	1398	4	US-09-925-065A-74751	Sequence 74751, A	C 360	44	1.4	470	8	US-10-674-124A-9871	Sequence 9871, Ap
C 288	45	1.4	1398	4	US-09-925-065A-74751	Sequence 74751, A	C 361	44	1.4	470	8	US-10-674-124A-9871	Sequence 9871, Ap
C 289	45	1.4	1792	4	US-09-925-065A-549910	Sequence 549910,	C 362	44	1.4	496	4	US-09-925-065A-876040	Sequence 876040,
C 290	45	1.4	1869	4	US-09-925-065A-549910	Sequence 549910,	C 363	44	1.4	496	4	US-09-925-065A-876042	Sequence 876042,
C 291	45	1.4	4513	6	US-10-264-237-2853	Sequence 2853, Ap	C 364	44	1.4	496	4	US-09-925-065A-876042	Sequence 876042,
C 292	45	1.4	6565	3	US-09-764-891-5739	Sequence 5739, Ap	C 365	44	1.4	496	4	US-09-925-065A-905372	Sequence 905372,
C 293	45	1.4	6565	3	US-09-764-891-5739	Sequence 5739, Ap	C 366	44	1.4	530	4	US-09-925-065A-297221	Sequence 297221,
C 294	45	1.4	17581	3	US-09-764-869-2170	Sequence 2170, Ap	C 367	44	1.4	530	4	US-09-925-065A-297222	Sequence 297222,
C 295	45	1.4	17581	3	US-10-911-504-2170	Sequence 2170, Ap	C 368	44	1.4	537	4	US-09-925-065A-448591	Sequence 448591,
C 296	45	1.4	17946	6	US-09-764-869-2312	Sequence 2312, Ap	C 369	44	1.4	537	4	US-09-925-065A-448592	Sequence 448592,
C 297	45	1.4	17946	6	US-10-091-504-2312	Sequence 2312, Ap	C 370	44	1.4	556	4	US-09-925-065A-109914	Sequence 109914,
C 298	45	1.4	17946	6	US-10-927-577-2312	Sequence 2312, Ap	C 371	44	1.4	557	4	US-09-925-065A-870384	Sequence 870384,
C 299	45	1.4	23694	5	US-10-087-192-1936	Sequence 1936, Ap	C 372	44	1.4	557	4	US-09-925-065A-870385	Sequence 870385,
C 300	45	1.4	23694	5	US-10-087-192-1936	Sequence 1936, Ap	C 373	44	1.4	565	4	US-09-925-065A-78244	Sequence 78244, A
C 301	45	1.4	3116	8	US-10-741-600-17761	Sequence 17761, A	C 374	44	1.4	565	4	US-10-027-633-127998	Sequence 127998,
C 302	45	1.4	32706	5	US-10-087-192-1660	Sequence 1660, Ap	C 375	44	1.4	565	6	US-10-027-633-127998	Sequence 127998,
C 303	45	1.4	32706	5	US-10-087-192-1750	Sequence 1750, Ap	C 376	44	1.4	577	4	US-09-925-065A-107681	Sequence 107681,
C 304	45	1.4	36534	6	US-10-285-351B-3	Sequence 3, Appl	C 377	44	1.4	577	4	US-09-925-065A-107682	Sequence 107682,
C 305	45	1.4	36534	7	US-10-240-425-1461	Sequence 1461, Ap	C 378	44	1.4	585	5	US-10-027-633-201634	Sequence 201634,
C 306	45	1.4	38684	9	US-10-893-315-136	Sequence 136, App	C 379	44	1.4	585	6	US-10-027-633-201634	Sequence 201634,
C 307	45	1.4	38684	9	US-10-893-315-136	Sequence 136, App	C 380	44	1.4	588	4	US-09-925-065A-129173	Sequence 129173,
C 308	45	1.4	39566	7	US-10-304-125-11	Sequence 11, Appl	C 381	44	1.4	588	4	US-09-925-065A-129173	Sequence 129173,
C 309	45	1.4	40090	7	US-10-644-021A-3	Sequence 3, Appl	C 382	44	1.4	589	5	US-10-027-633-197271	Sequence 197271,
C 310	45	1.4	65277	9	US-10-461-862-152	Sequence 152, App	C 383	44	1.4	589	5	US-10-027-633-197272	Sequence 197272,
C 311	45	1.4	73995	5	US-10-087-192-208	Sequence 208, App	C 384	44	1.4	589	6	US-10-027-633-197271	Sequence 197271,
C 312	45	1.4	83900	9	US-10-722-936-3	Sequence 3, Appl	C 385	44	1.4	589	6	US-10-027-633-197272	Sequence 197272,
C 313	45	1.4	103574	9	US-10-756-149-2307	Sequence 2307, Ap	C 386	44	1.4	590	5	US-10-027-633-94737	Sequence 94737, A
C 314	45	1.4	107543	7	US-10-322-281-706	Sequence 706, App	C 387	44	1.4	590	5	US-10-027-633-318558	Sequence 318558,
C 315	45	1.4	107745	7	US-10-322-281-268	Sequence 268, App	C 388	44	1.4	590	6	US-10-027-633-94737	Sequence 94737, A

C 389 44 1.4 590 6 US-10-027-632-318528 Sequence 318528,  
C 390 44 1.4 591 4 US-09-925-065A-936386 Sequence 936386,  
C 391 44 1.4 593 4 US-09-925-065A-237003 Sequence 237003,  
C 392 44 1.4 593 4 US-09-925-065A-237005 Sequence 237005,  
C 393 44 1.4 594 4 US-09-925-065A-685243 Sequence 685243,  
C 394 44 1.4 595 4 US-09-925-065A-902996 Sequence 902996,  
C 395 44 1.4 600 9 US-10-956-157-73499 Sequence 73499, Ap  
C 396 44 1.4 611 7 US-10-276-774-491 Sequence 491, App  
C 397 44 1.4 612 4 US-09-925-065A-21972 Sequence 21972, A  
C 398 44 1.4 614 5 US-10-027-632-260590 Sequence 260590,  
C 399 44 1.4 614 6 US-10-027-632-260590 Sequence 260590,  
C 400 44 1.4 618 4 US-09-925-065A-697941 Sequence 697941,  
C 401 44 1.4 618 4 US-09-925-065A-697942 Sequence 697942,  
C 402 44 1.4 618 4 US-09-925-065A-697943 Sequence 697943,  
C 403 44 1.4 619 4 US-09-925-065A-789282 Sequence 789282,  
C 404 44 1.4 619 4 US-09-925-065A-789283 Sequence 789283,  
C 405 44 1.4 619 4 US-09-925-065A-789284 Sequence 789284,  
C 406 44 1.4 619 4 US-09-925-065A-847245 Sequence 847245,  
C 407 44 1.4 619 4 US-09-925-065A-951338 Sequence 951338,  
C 408 44 1.4 623 4 US-09-925-065A-521713 Sequence 521713,  
C 409 44 1.4 627 4 US-09-925-065A-545865 Sequence 545865,  
C 410 44 1.4 627 4 US-09-925-065A-545866 Sequence 545866,  
C 411 44 1.4 627 4 US-09-925-065A-545867 Sequence 545867,  
C 412 44 1.4 627 4 US-09-925-065A-923432 Sequence 923432,  
C 413 44 1.4 629 4 US-09-925-065A-821309 Sequence 821309,  
C 414 44 1.4 637 5 US-10-027-632-187929 Sequence 187929,  
C 415 44 1.4 637 5 US-10-027-632-187929 Sequence 187929,  
C 416 44 1.4 643 4 US-09-925-065A-692017 Sequence 692017,  
C 417 44 1.4 659 4 US-09-925-065A-741517 Sequence 741517,  
C 418 44 1.4 659 4 US-09-925-065A-741518 Sequence 741518,  
C 419 44 1.4 668 4 US-09-925-065A-127024 Sequence 127024,  
C 420 44 1.4 673 4 US-09-925-065A-32538 Sequence 32538, A  
C 421 44 1.4 698 5 US-10-027-632-252317 Sequence 252317,  
C 422 44 1.4 698 5 US-10-027-632-252317 Sequence 252317,  
C 423 44 1.4 738 5 US-10-027-632-139602 Sequence 139602,  
C 424 44 1.4 738 5 US-10-027-632-139602 Sequence 139602,  
C 425 44 1.4 738 5 US-10-027-632-139602 Sequence 139602,  
C 426 44 1.4 748 5 US-10-027-632-172816 Sequence 172816,  
C 427 44 1.4 748 5 US-10-027-632-146061 Sequence 146061,  
C 428 44 1.4 748 5 US-10-027-632-146061 Sequence 146061,  
C 429 44 1.4 751 5 US-10-027-632-14193 Sequence 14193, A  
C 430 44 1.4 751 5 US-10-027-632-14193 Sequence 14193, A  
C 431 44 1.4 755 5 US-10-027-632-27592 Sequence 27592, A  
C 432 44 1.4 755 5 US-10-027-632-27592 Sequence 27592, A  
C 433 44 1.4 755 5 US-10-027-632-27592 Sequence 27592, A  
C 434 44 1.4 755 5 US-10-027-632-27592 Sequence 27592, A  
C 435 44 1.4 756 5 US-10-027-632-24620 Sequence 24620, A  
C 436 44 1.4 756 5 US-10-027-632-149995 Sequence 149995, A  
C 437 44 1.4 756 5 US-10-027-632-149995 Sequence 149995, A  
C 438 44 1.4 784 4 US-09-925-065A-927249 Sequence 927249,  
C 439 44 1.4 810 5 US-10-027-632-154967 Sequence 154967,  
C 440 44 1.4 810 5 US-10-027-632-154967 Sequence 154967,  
C 441 44 1.4 827 5 US-10-027-632-148684 Sequence 148684,  
C 442 44 1.4 827 5 US-10-027-632-148684 Sequence 148684,  
C 443 44 1.4 827 5 US-10-027-632-148684 Sequence 148684,  
C 444 44 1.4 827 5 US-10-027-632-148684 Sequence 148684,  
C 445 44 1.4 827 5 US-10-027-632-148684 Sequence 148684,  
C 446 44 1.4 873 5 US-10-027-632-254541 Sequence 254541,  
C 447 44 1.4 873 5 US-10-027-632-254541 Sequence 254541,  
C 448 44 1.4 956 8 US-10-170-097-56 Sequence 56, Appl  
C 449 44 1.4 956 8 US-10-170-097-56 Sequence 56, Appl  
C 450 44 1.4 1005 4 US-09-925-065A-726951 Sequence 726951,  
C 451 44 1.4 1005 4 US-09-925-065A-726952 Sequence 726952,  
C 452 44 1.4 1051 5 US-10-027-632-31546 Sequence 31546, A  
C 453 44 1.4 1051 5 US-10-027-632-31547 Sequence 31547, A  
C 454 44 1.4 1051 5 US-10-027-632-31548 Sequence 31548, A  
C 455 44 1.4 1051 6 US-10-027-632-31546 Sequence 31546, A  
C 456 44 1.4 1051 6 US-10-027-632-31547 Sequence 31547, A  
C 457 44 1.4 1051 6 US-10-027-632-31548 Sequence 31548, A  
C 458 44 1.4 1068 4 US-09-925-065A-716304 Sequence 716304, A  
C 459 44 1.4 1068 4 US-09-925-065A-716305 Sequence 716305, A  
C 460 44 1.4 1068 4 US-09-925-065A-716306 Sequence 716306, A  
C 461 44 1.4 1068 4 US-09-925-065A-716307 Sequence 716307, A

462 44 1.4 1068 4 US-09-925-065A-716308 Sequence 716308,  
C 463 44 1.4 1103 5 US-10-027-632-257954 Sequence 257954,  
C 464 44 1.4 1103 5 US-10-027-632-257955 Sequence 257955,  
C 465 44 1.4 1103 6 US-10-027-632-257954 Sequence 257954,  
C 466 44 1.4 1103 6 US-10-027-632-257955 Sequence 257955,  
C 467 44 1.4 1122 4 US-09-925-065A-551840 Sequence 551840,  
C 468 44 1.4 1240 5 US-10-027-632-263693 Sequence 263693,  
C 469 44 1.4 1240 5 US-10-027-632-263693 Sequence 263693,  
C 470 44 1.4 1242 4 US-09-925-065A-717995 Sequence 717995,  
C 471 44 1.4 1330 4 US-09-925-065A-583322 Sequence 58322, A  
C 472 44 1.4 1330 4 US-09-925-065A-583323 Sequence 58323, A  
C 473 44 1.4 1330 4 US-09-925-065A-583434 Sequence 58344, A  
C 474 44 1.4 1390 4 US-09-925-065A-167962 Sequence 167962, A  
C 475 44 1.4 1427 4 US-09-925-065A-177993 Sequence 177993, A  
C 476 44 1.4 1734 5 US-10-027-632-100291 Sequence 100291, A  
C 477 44 1.4 1734 6 US-10-027-632-100291 Sequence 100291, A  
C 478 44 1.4 1753 4 US-09-925-065A-714611 Sequence 714611, A  
C 479 44 1.4 1753 4 US-09-925-065A-714612 Sequence 714612, A  
C 480 44 1.4 1753 4 US-09-925-065A-714613 Sequence 714613, A  
C 481 44 1.4 1753 4 US-09-925-065A-714614 Sequence 714614, A  
C 482 44 1.4 1900 5 US-10-027-632-263112 Sequence 263112, A  
C 483 44 1.4 1900 6 US-10-027-632-263112 Sequence 263112, A  
C 484 44 1.4 1970 4 US-09-925-065A-121388 Sequence 121388, A  
C 485 44 1.4 2078 4 US-09-925-065A-167966 Sequence 167966, A  
C 486 44 1.4 2184 4 US-09-925-065A-175885 Sequence 175885, A  
C 487 44 1.4 2184 4 US-09-925-065A-175886 Sequence 175886, A  
C 488 44 1.4 2184 4 US-09-925-065A-175887 Sequence 175887, A  
C 489 44 1.4 2184 4 US-09-925-065A-175888 Sequence 175888, A  
C 490 44 1.4 2725 4 US-09-925-065A-696477 Sequence 696477, A  
C 491 44 1.4 2845 9 US-10-956-157-2234 Sequence 2234, Ap  
C 492 44 1.4 3488 4 US-09-925-065A-84868 Sequence 84868, A  
C 493 44 1.4 5307 5 US-10-074-045-59 Sequence 59, Appl  
C 494 44 1.4 5307 5 US-10-074-045-59 Sequence 59, Appl  
C 495 44 1.4 6467 6 US-10-172-118-15 Sequence 61, Appl  
C 496 44 1.4 6467 7 US-10-342-887-15 Sequence 15, Appl  
C 497 44 1.4 6467 9 US-10-756-149-2274 Sequence 2274, Ap  
C 498 44 1.4 7720 3 US-09-954-456-946 Sequence 946, Ap  
C 499 44 1.4 7720 3 US-09-954-456-1589 Sequence 1589, Ap  
C 500 44 1.4 7720 3 US-09-968-007A-125 Sequence 125, Ap

## ALIGNMENTS

RESULT 1  
US-10-450-763-29532  
; Sequence 29532, Application US/10450763  
; Publication No. US2005019675A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; PRIOR FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; NUMBER OF SEQ ID NOS: 2000-08-23  
; SOFTWARE: Custom  
; SEQ ID NO 29532  
; LENGTH: 850  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: SIMILAR  
; LOCATION: (241)..(699)  
; OTHER INFORMATION: 30% homologous to Leishmania major L8453.1, accession number  
US-10-450-763-29532

Query Match 24.0%; Score 748; DB 9; Length 850;  
Best Local Similarity 99.8%; Pred. No. 0;  
Matches 848; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 716 CGGTACCGCCGAGCGATCAAGCAATGAAGAGCGTTGTCGCGCGCCGAGCGCG 775  
DB 1 CGGTACCGCCGAGCGATCAAGCAATGAAGAGCGTTGTCGCGCGCCGAGCGCG 60  
QY 776 GGAATGGGGGTTAGCCATCTCTGCGCGCTGAGGGGGAGGCTTAAACGGCGCGCGCG 835  
DB 61 GGAATGGGGGTTAGCCATCTCTGCGCGCTGAGGGGGAGGCTTAAACGGCGCGCGCG 120  
QY 836 GGCACCGCCGAGCCCAACCGAGTGGCGAGAGAGAGTCAAGAGCGCTCTGAGCGAGCT 895  
DB 121 GGCACCGCCGAGCCCAACCGAGTGGCGAGAGAGAGTCAAGAGCGCTCTGAGCGAGCT 180  
QY 896 CAACAAGAGAGAGTGGCTGCTCAACAACCTGAGTGGAGCGTGGTGGCTCGCGAGACTC 955  
DB 181 CAACAAGAGAGAGTGGCTGCTCAACAACCTGAGTGGAGCGTGGTGGCTCGCGAGACTC 240  
QY 956 GGAAGAACCTGGCGAGAGAGTGGAGAAAGAGCGCGCGAGAGAGCGAGAGAGTGGCGGTC 1015  
DB 241 GGAAGAACCTGGCGAGAGAGTGGAGAAAGAGCGCGCGAGAGAGCGAGAGAGTGGCGGTC 300  
QY 1016 CACCTGCGCGCGAGTGAAGTGGTGGCGAGACCGGGCGCTGGCGCGAGCGAGCGCG 1075  
DB 301 CACCTGCGCGCGAGTGAAGTGGTGGCGAGACCGGGCGCTGGCGCGAGCGAGCGCG 360  
QY 1076 CGAGTTGAGAGCGAGTGGAGTGGCTTCTGCGAGCTGGAGAGCTTGGAGAGCGAGCAT 1135  
DB 361 CGAGTTGAGAGCGAGTGGAGTGGCTTCTGCGAGCTGGAGAGCTTGGAGAGCGAGCAT 420  
QY 1136 GCGAGCGCTGGTGAAGTGGAGCGCGCGCTTCCGCTGGAGACCGCGCGCGAGCGCTGGT 1195  
DB 421 GCGAGCGCTGGTGAAGTGGAGCGCGCGCTTCCGCTGGAGACCGCGCGCGAGCGCTGGT 480  
QY 1196 GCGAGAGAGTGGAGTGGAGCGCTCTCCGAGCGTGGAGCGCGCGCTGGAGAGCGAG 1255  
DB 481 GCGAGAGAGTGGAGTGGAGCGCTCTCCGAGCGTGGAGCGCGCGCTGGAGAGCGAG 540  
QY 1256 CCGTGGCGCTGAGAGCGAGCGAGCGAGCTTGAAGTGGAGAGCTGGAGAGCGAG 1315  
DB 541 CCGTGGCGCTGAGAGCGAGCGAGCGAGCTTGAAGTGGAGAGCTGGAGAGCGAG 600  
QY 1316 GGTCTTCAAGTGGAGAGTGAATGAACAATGAAGAGTGAAGTGAAGTGAAGTGAAGTGA 1375  
DB 601 GGTCTTCAAGTGGAGAGTGAATGAACAATGAAGAGTGAAGTGAAGTGAAGTGAAGTGA 660  
QY 1376 GACCGTGAAGCGCGAGCGAGCGAGCGAGCTCTGCAAGTGAAGCGCGAGCGCG 1435  
DB 661 GACCGTGAAGCGCGAGCGAGCGAGCGAGCGAGCTCTGCAAGTGAAGCGCGAGCGCG 720  
QY 1436 CTCCTGAGTGGTCTCTTGAAGAGCGAGCGAGAGTGGAGAGCGAGAGCGCTGGAG 1495  
DB 721 CTCCTGAGTGGTCTCTTGAAGAGCGAGCGAGAGTGGAGAGCGAGAGCGCTGGAG 780  
QY 1496 CGCGATCTTTTGGAGCGCGTGTCTGCGAGCTTGGAGCGCTTGGAGCGTGGAGCGA 1555  
DB 781 CGCGATCTTTTGGAGCGCGTGTCTGCGAGCTTGGAGCGCTTGGAGCGTGGAGCGA 840  
QY 1556 GCTGAGCTGA 1565  
DB 841 GCTGAGCTGA 850

RESULT 2  
US-10-029-386-20699/c  
; Sequence 20699; Application US/10029386  
; Publication No. US20030194704A1  
; GENERAL INFORMATION:  
; APPLICANT: Penn, Sharon G.  
; APPLICANT: Rank, David R.  
; APPLICANT: Hanzel, David K.

; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR  
; TITLE OF INVENTION: EXPRESSION ANALYSIS TWO  
; FILE REFERENCE: AEMICA-X-2  
; CURRENT APPLICATION NUMBER: US/10/029,386  
; CURRENT FILING DATE: 2001-12-20  
; NUMBER OF SEQ ID NOS: 34288  
; SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
; SEQ ID NO 20699  
; LENGTH: 708  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; OTHER INFORMATION: MAP TO AC010615.5  
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 0.56  
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2  
; OTHER INFORMATION: EXPRESSED IN HEPA, SIGNAL = 1.3  
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1  
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5  
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3  
; OTHER INFORMATION: EST HUMAN HIT: AM302149.1, EVALUR 0.00e+00  
; OTHER INFORMATION: SWISSPROT HIT: P31735, EVALUR 5.50e+00  
US-10-029-386-20699

Query Match 21.0%; Score 657; DB 6; Length 708;  
Best Local Similarity 99.9%; Pred. No. 0;  
Matches 707; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 858 ATGGGAGAGAGAGTGAAGAGCGTCTGAGAGCGAGCTCAACAAGAGAGTGGCTGAC 917  
DB 708 ATGGGAGAGAGAGTGAAGAGCGTCTGAGAGCGAGCTCAACAAGAGAGTGGCTGAC 649  
QY 918 CACCACTGGTGTCTACACGTGGTGGCTGGAGAGTGGAGAGTGGAGAGTGGAGAGTGG 977  
DB 648 CACCACTGGTGTCTACACGTGGTGGCTGGAGAGTGGAGAGTGGAGAGTGGAGAGTGG 589  
QY 978 CAAAAGAGCGCGAGAGAGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 1037  
DB 588 CAAAAGAGCGCGAGAGAGAGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAG 529  
QY 1038 GTGCTGGCGAGACCGGGGCTGGCGCGCGAGCGAGCGAGTGGAGTGGAGTGGAGTGG 1097  
DB 528 GTGCTGGCGAGACCGGGGCTGGCGCGCGAGCGAGCGAGTGGAGTGGAGTGGAGTGG 469  
QY 1098 GCTTCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 1157  
DB 468 GCTTCTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGGAGTGG 409  
QY 1158 GCGGAGTTCGCGTCAAGCGCGCGAGCGAGCGAGTGGAGTGGAGTGGAGTGGAGTGG 1217  
DB 408 GCGGAGTTCGCGTCAAGCGCGCGAGCGAGCGAGTGGAGTGGAGTGGAGTGGAGTGG 349  
QY 1218 TCTTCTGGAGTGGAGCGCGCGCGTGAAGACCGCGAGTGGAGTGGAGTGGAGTGGAG 1277  
DB 348 TCTTCTGGAGTGGAGCGCGCGCGTGAAGACCGCGAGTGGAGTGGAGTGGAGTGGAG 289  
QY 1278 GACTTGAAGTGGAGCGCGCGCGTGAAGACCGCGAGTGGAGTGGAGTGGAGTGGAGTGG 1337  
DB 288 GACTTGAAGTGGAGCGCGCGCGTGAAGACCGCGAGTGGAGTGGAGTGGAGTGGAGTGG 229  
QY 1338 ATGCAACAATGAGATGAAGTCAACGTGCGCGCGTGGAGCGTGGAGCGAGCGAGCG 1397  
DB 228 ATGCAACAATGAGATGAAGTCAACGTGCGCGCGTGGAGCGTGGAGCGAGCGAGCGAG 169  
QY 1398 GCGGAGCGAGAGTCTCTTCAAGTGAAGCGCGCGCGTCTCTGAGTGGTGGTGGAG 1457  
DB 168 GCGGAGCGAGAGTCTCTTCAAGTGAAGCGCGCGCGTCTCTGAGTGGTGGTGGAG 109  
QY 1458 GAGCGCGAGAGAGTGGAGAGCGCGAGAGAGCGTGGAGCGAGTGGAGCGAGTGGAG 1517  
DB 108 GAGCGCGAGAGAGTGGAGAGCGCGAGAGAGCGTGGAGCGAGTGGAGCGAGTGGAG 49  
QY 1518 CTGCTGGAGAGTGGAGCGTGGAGCGTGGAGCGTGGAGCGAGTGGAGCGAGTGGAG 1565

Db 48 CTGCTGGCGGCTGTGTGGCCCTAGCCGCTGTGTGGGAGAGCTGACTGA 1

RESULT 3  
US-10-029-386-6988/c  
Sequence 6988, Application US/10029386  
Publication No. US20030194704A1  
GENERAL INFORMATION:  
APPLICANT: Penn, Sharon G.  
APPLICANT: Rank, David R.  
APPLICANT: Hanzel, David K.  
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR G  
FILE REFERENCE: ABOMICA-X-2  
CURRENT FILING DATE: 2001-12-20  
NUMBER OF SEQ ID NOS: 34288  
SOFTWARE: Annomax Sequence Listing Engine vers. 1.1  
SEQ ID NO 6988  
LENGTH: 524  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: MAP TO ACO10615.5  
OTHER INFORMATION: EXPRESSED IN HEILA, SIGNAL = 0.56  
OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.2  
OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.3  
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1  
OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 1.5  
OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.3  
OTHER INFORMATION: EST HUMAN HIT: AW302149.1, EVALU0.00e+00  
OTHER INFORMATION: SWISSPROT HIT: 006805, EVALU0.130e-01  
OTHER INFORMATION: NT HIT: g14507086, EVALU0.7.80e+00  
US-10-029-386-6988

Query Match 16.8%; Score 524; DB 6; Length 524;  
Best Local Similarity 100.0%; Pred. No. 6.5e-265; Indels 0; Gaps 0;

Matches 524; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 484 CTCTGAGCCTGCTTGGCCCGGAGTTGGACCCACGAGAGATGGGACCGACCTCAGC 543  
DB 524 CTCTGAGCCTGCTTGGCCCGGAGTTGGACCCACGAGAGATGGGACCGACCTCAGC 465  
QY 544 TTTCGAGGAGGACCGTGTGAGGCGGCGGTGCGAGAGACGAGTGTGACTGGAGT 603  
DB 464 TTTCGAGGAGGACCGTGTGAGGCGGCGGTGCGAGAGACGAGTGTGACTGGAGT 405  
QY 604 GCGCCTGGGAGAGATGAGACGAGGAGCGGAGGACCGCTAACGGGGCTCCCTCTGCGCGC 663  
DB 404 GCGCCTGGGAGAGATGAGACGAGGAGCGGAGGACCGCTAACGGGGCTCCCTCTGCGCGC 345  
QY 664 CCGCTCGCAGAGGCGACGTCGAGGGTCCCGGCGGGGCTCCCTGAGCGTGGCGTACG 723  
DB 344 CCGCTCGCAGAGGCGACGTCGAGGGTCCCGGCGGGGCTCCCTGAGCGTGGCGTACG 285  
QY 724 CCGAGGAGATGACGAGCAATGAAGAGGTTGTCGCGCGCGCGCCAAAGCCGGGATGGG 783  
DB 284 CCGAGGAGATGACGAGCAATGAAGAGGTTGTCGCGCGCGCGCCAAAGCCGGGATGGG 225  
QY 784 GTTAGCCATCTCTGCGCGCTGAGGAGGAGGCTTAAACGAGCGCGCGCGCCAGC 843  
DB 224 GTTAGCCATCTCTGCGCGCTGAGGAGGAGGCTTAAACGAGCGCGCGCGCCAGC 165  
QY 844 CCGAGCCACCGGAGATGGCGAGGAGAGTGCAGAGCGCTGTGACCGGCTCAACAAG 903  
DB 164 CCGAGCCACCGGAGATGGCGAGGAGAGTGCAGAGCGCTGTGACCGGCTCAACAAG 105  
QY 904 CGACTGCGTCAACCACTGTGTGCTGACCGTGTGTGCTGTGAGGAGCTGTGACAGAC 963  
DB 104 CGACTGCGTCAACCACTGTGTGCTGACCGTGTGTGCTGTGAGGAGCTGTGACAGAC 45  
QY 964 TCGCGAGAGCTGCAAAAGACCGCGCAGAGGCGGAGAGCTG 1007

Db 44 TCGCGGAGAGCTGCAAAAAGACCGCCGAGAAAGCGGACGAGAGCTG 1

RESULT 4  
US-09-925-065A-740956/c  
Sequence 740956, Application US/09925065A  
Publication No. US20050228172A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single  
FILE REFERENCE: 108827.135  
CURRENT APPLICATION NUMBER: US/09/925,065A  
CURRENT FILING DATE: 2001-08-08  
PRIOR APPLICATION NUMBER: US 60/243,096  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: US 60/252,147  
PRIOR FILING DATE: 2000-11-20  
PRIOR APPLICATION NUMBER: US 60/250,092  
PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: US 60/261,766  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/289,846  
NUMBER OF SEQ ID NOS: 957086  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 740956  
LENGTH: 560  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-925-065A-740956

Query Match 16.3%; Score 509; DB 4; Length 560;  
Best Local Similarity 99.8%; Pred. No. 5.3e-257; Indels 0; Gaps 0;

Matches 559; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 199 CATCAGAGCAAAACTCCGCGAGGAGCTGCGCCCTTTTAACTGGGCTTCACTGTTCC 258  
DB 560 CATCAGAGCAAAACTCCGCGAGGAGCTGCGCCCTTTTAACTGGGCTTCACTGTTCC 501  
QY 259 CATCCGTAATAATGAAACGGGTGATCTCCGAGCGCTAACATTCAGAACTCGATGG 318  
DB 500 CATCCGTAATAATGAAACGGGTGATCTCCGAGCGCTAACATTCAGAACTCGATGG 441  
QY 319 GCGAAGGAGGAGGATGGGCAACCAACAGTAACTCCCGGCTGAGGCGCCCTA 378  
DB 440 GCGAAGGAGGAGGATGGGCAACCAACAGTAACTCCCGGCTGAGGCGCCCTA 381  
QY 379 CCACTGATCCAGAGGAGTGGAGCTCCGCGCGGAGCAGACGAGGATGGGCGTCTAAGAA 438  
DB 380 CCACTGATCCAGAGGAGTGGAGCTCCGCGCGGAGCAGACGAGGATGGGCGTCTAAGAA 321  
QY 439 ACCCTACCCGCGCTCTTGGAGCGCTTAAGCGGAGCGGCGCTCTGACGCTGCTT 498  
DB 320 ACCCTACCCGCGCTCTTGGAGCGCTTAAGCGGAGCGGCGCTCTGACGCTGCTT 261  
QY 499 GCCCGGAGTTGGACCCACGAGGAGTGGGAGCGGACCCCTAGGCTTGGAGGAGCCAC 558  
DB 260 GCCCGGAGTTGGACCCACGAGGAGTGGGAGCGGACCCCTAGGCTTGGAGGAGCCAC 201  
QY 559 CGTGAAGCCAGAGGCGGTGCAAGACAAGACGTGTGACTCGAGTGTGCGCTGGAGAGAT 618  
DB 200 CGTGAAGCCAGAGGCGGTGCAAGACAAGACGTGTGACTCGAGTGTGCGCTGGAGAGAT 141  
QY 619 GGAAGAGGAGCGGAGACCGCTTAAACGGGCTTCTGCGCGCGCGCTTCCGAGAGGCG 678  
DB 140 GGAAGAGGAGCGGAGACCGCTTAAACGGGCTTCTGCGCGCGCGCTTCCGAGAGGCG 81  
QY 679 CAAGTCAGAGGCTCCGCGCGGCTCCGAGAGCTTGGAGGAGTGGAGGCGGAGGAGTCAAG 738  
DB 80 CAAGTCAGAGGCTCCGCGCGGCTCCGAGAGCTTGGAGGAGTGGAGGCGGAGGAGTCAAG 21  
QY 739 ACCATGAAGAGCTTCTGTC 758



Db 20 ACCATGAAGAGCGTTCGTGC 1

## RESULT 5

US-09-925-065A-769508/c  
; Sequence 769508, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 769508  
; LENGTH: 557  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-769508

Query Match 16.2%; Score 506; DB 4; Length 557;  
Best Local Similarity 99.8%; Pred. No. 2e-255;  
Matches 556; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 202 CACAGGCAAAACTCCGCGGAGCGCTGGCCGCTTTTTCCTGAGGCTCAGTTTCCCAT 261  
DB 557 CACAGGCAAAACTCCGCGGAGCGCTGGCCGCTTTTTCCTGAGGCTCAGTTTCCCAT 498  
QY 262 CCGTAATAATAGAACGGGTTTGATCTCCGAGCGCTAACATTCCAGAACTCGATGGGCG 321  
DB 497 CCGTAATAATAGAACGGGTTTGATCTCCGAGCGCTAACATTCCAGAACTCGATGGGCG 438  
QY 322 AAGGGAGGAGGAGATGGGCGCACTGACCTCCCGCGTGGAGCCCGGCTTACCA 381  
DB 437 AAGGGAGGAGGAGATGGGCGCACTGACCTCCCGCGTGGAGCCCGGCTTACCA 378  
QY 382 CTGATCCAGGGGGTGGCACTCCGCGCGGAGCGAGCGGGGTGGGCGGGCTTAAAGAAAC 441  
DB 377 CTGATCCAGGGGGTGGCACTCCGCGCGGAGCGAGCGGGGTGGGCGGGCTTAAAGAAAC 318  
QY 442 CTACCCGCGCGCTTGGGAGCGCTTAAGCGGAGCGCGGCTCTGACGCTTGTGCG 501  
DB 317 CTACCCGCGCGCTTGGGAGCGCTTAAGCGGAGCGCGGCTCTGACGCTTGTGCG 258  
QY 502 CCGGAGTTGGCACTCCGAGGATGGGAGCCGACCTTACGCTTGCAGAGGAGCCAGCT 561  
DB 257 CCGGAGTTGGCACTCCGAGGATGGGAGCCGACCTTACGCTTGCAGAGGAGCCAGCT 198  
QY 562 GAGAGCCAGGCGGCTGACAGACAGACGCTGAGCTCGAGGTCGCGTGGGAGATGGA 621  
DB 197 GAGAGCCAGGCGGCTGACAGACAGACGCTGAGCTCGAGGTCGCGTGGGAGATGGA 138  
QY 622 CGAGGAGCGGAGGAGCGCTTAAAGGAGCTCCCTTTCGCGCGCCCGCTCGCAGAGGCGAC 681  
DB 137 CGAGGAGCGGAGGAGCGCTTAAAGGAGCTCCCTTTCGCGCGCCCGCTCGCAGAGGCGAC 78  
QY 682 GTCCAGAGGTCCTCGGCGGCGCTCCGTGGAAGTTGGCGGTAGCCCGGAGAGTCAAGGAGC 741  
DB 77 GTCCAGAGGTCCTCGGCGGCGCTCCGTGGAAGTTGGCGGTAGCCCGGAGAGTCAAGGAGC 18

QY 742 ATGAAGAGCGTTCGTGC 758  
DB 17 ATGAAGAGCGTTCGTGC 1

## RESULT 6

US-09-925-065A-736351/c  
; Sequence 736351, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 736351  
; LENGTH: 556  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-736351

Query Match 16.2%; Score 505; DB 4; Length 556;  
Best Local Similarity 99.8%; Pred. No. 6.8e-255;  
Matches 555; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 203 AAGGCAAAACTCCGCGGAGCGCTGGCCGCTTTTTCCTGAGGCTCAGTTTCCCAT 262  
DB 556 AAGGCAAAACTCCGCGGAGCGCTGGCCGCTTTTTCCTGAGGCTCAGTTTCCCAT 497  
QY 263 CCGTAATAATAGAACGGGTTTGATCTCCGAGCGCTAACATTCCAGAACTCGATGGGCG 322  
DB 496 CCGTAATAATAGAACGGGTTTGATCTCCGAGCGCTAACATTCCAGAACTCGATGGGCG 437  
QY 322 AAGGGAGGAGGAGATGGGCGCACTGACCTCCCGCGTGGAGCCCGGCTTACCA 382  
DB 436 AAGGGAGGAGGAGATGGGCGCACTGACCTCCCGCGTGGAGCCCGGCTTACCA 377  
QY 382 TGATCCAGGGGGTGGCACTCCGCGCGGAGCGAGCGGGGTGGGCGGGCTTAAAGAAAC 442  
DB 376 TGATCCAGGGGGTGGCACTCCGCGCGGAGCGAGCGGGGTGGGCGGGCTTAAAGAAAC 317  
QY 443 TACCCGCGCGCTTGGGAGCGCTTAAGCGGAGCGCGGCTCTGACGCTTGTGCG 502  
DB 316 TACCCGCGCGCTTGGGAGCGCTTAAGCGGAGCGCGGCTCTGACGCTTGTGCG 257  
QY 503 CCGGAGTTGGCACTCCGAGGATGGGAGCCGACCTTACGCTTGCAGAGGAGCCAGCT 562  
DB 256 CCGGAGTTGGCACTCCGAGGATGGGAGCCGACCTTACGCTTGCAGAGGAGCCAGCT 197  
QY 562 GAGGCGAGGCGGTCAGAGACAGACGCTGAGCTCGAGGTCGCGTGGGAGATGGA 622  
DB 196 GAGGCGAGGCGGTCAGAGACAGACGCTGAGCTCGAGGTCGCGTGGGAGATGGA 137  
QY 622 GAGGAGCGGAGGAGCGCTTAAAGGAGCTCCCTTTCGCGCGCCCGCTCGCAGAGGCGAC 682  
DB 136 GAGGAGCGGAGGAGCGCTTAAAGGAGCTCCCTTTCGCGCGCCCGCTCGCAGAGGCGAC 77  
QY 682 TCCAGAGGTCCTCGGCGGCGCTCCGTGGAAGTTGGCGGTAGCCCGGAGAGTCAAGGAGC 742  
DB 76 TCCAGAGGTCCTCGGCGGCGCTCCGTGGAAGTTGGCGGTAGCCCGGAGAGTCAAGGAGC 17

QY 743 TGAAGAGCGTTCGTGC 758  
Db 16 TGAAGAGCGTTCGTGC 1

## RESULT 7

US-09-925-065A-737120  
; Sequence 737120, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108927.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 737120  
; LENGTH: 555  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-737120

Query Match 16.1%; Score 504; DB 4; Length 555;  
Best Local Similarity 9.8%; Pred. No. 2.3e-254;  
Matches 554; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 204 CAGGCAAAAACCTCCGGGAGACCTGGCCCGCTTTTAACTGGGCTTCAGTTTCCCATTC 263  
Db 1 CAGGCAAAAACCTCCGGGAGACCTGGCCCGCTTTTAACTGGGCTTCAGTTTCCCATTC 60  
QY 264 GTPAAATAGAGGGGTTGATCTCCGAGGCTTAACATTCAGAACTCGAATGGGGCGAA 323  
Db 61 GTPAAATAGAGGGGTTGATCTCCGAGGCTTAACATTCAGAACTCGAATGGGGCGAA 120  
QY 324 GGGGAGGAGGATGGGCAACCCACAGTGAACCTCCCGCTGGAGCCCGCTACACT 383  
Db 121 GGGGAGGAGGATGGGCAACCCACAGTGAACCTCCCGCTGGAGCCCGCTACACT 180  
QY 384 GATCCAGGGGATGGAGCTCCCGCCGGAGAGAGCGGGGATGGGCTCTAGAAAACCT 443  
Db 181 GATCCAGGGGATGGAGCTCCCGCCGGAGAGAGCGGGGATGGGCTCTAGAAAACCT 240  
QY 444 ACCCGGCGCCCTTGGCAGGCTTAAGGAGGAGCGGGCTTGAGGCTGTTGCCCC 503  
Db 241 ACCCGGCGCCCTTGGCAGGCTTAAGGAGGAGCGGGCTTGAGGCTGTTGCCCC 300  
QY 504 GGAGTTGGACCCACGAGGATGGGACCGCACCTTCAGCTTTCAGGAGGCCACCGTGG 563  
Db 301 GGAGTTGGACCCACGAGGATGGGACCGCACCTTCAGCTTTCAGGAGGCCACCGTGG 360  
QY 564 AGGCCAGGGGCTGTCAGAGACAGCTGTGACTCGAGTGGCCCTGGGGAGGATGAGC 623  
Db 361 AGGCCAGGGGCTGTCAGAGACAGCTGTGACTCGAGTGGCCCTGGGGAGGATGAGC 420  
QY 624 AGGGAAGGGGGAGCGCTAAGCGGGGCTCCTCTGGGCGCCCGCTCGCAGAGGCGGACGT 683  
Db 421 AGGGAAGGGGGAGCGCTAAGCGGGGCTCCTCTGGGCGCCCGCTCGCAGAGGCGGACGT 480  
QY 684 CGAGGATCCGGGCGGCTCCGTGAGAGTTGGGCGTAAAGCCGAGCGAGTCAAGCAAT 743

Db 481 CGAGGATCCGGGCGGCTCCGTGAGAGTTGGCGGTAGCGCGAGCGAGTCAAGCAAT 540  
QY 744 GAAGAGCGTTCGTGC 758  
Db 541 GAAGAGCGTTCGTGC 555

## RESULT 8

US-10-450-763-8312/C  
; Sequence 8312, Application US/10450763  
; Publication No. US20050196754A1  
; GENERAL INFORMATION:  
; APPLICANT: Hyseq, Inc  
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
; FILE REFERENCE: 790CIP3/US  
; CURRENT APPLICATION NUMBER: US/10/450,763  
; CURRENT FILING DATE: 2003-06-11  
; PRIOR APPLICATION NUMBER: PCT/US01/08631  
; PRIOR FILING DATE: 2001-03-30  
; PRIOR APPLICATION NUMBER: 09/540,217  
; PRIOR FILING DATE: 2000-03-31  
; PRIOR APPLICATION NUMBER: 09/649,167  
; PRIOR FILING DATE: 2000-08-23  
; NUMBER OF SEQ ID NOS: 60736  
; SOFTWARE: Custom  
; SEQ ID NO 8312  
; LENGTH: 1349  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SIMILAR  
; LOCATION: (17)..(1054)  
; OTHER INFORMATION: 30\* homologous to Streptomyces coelicolor A3(2) putative  
; OTHER INFORMATION: translation initiation factor IF-2(fragment), accession number  
; OTHER INFORMATION: AL35913, Smith-Waterman Score=196.  
US-10-450-763-8312

Query Match 7.7%; Score 241; DB 9; Length 1349;  
Best Local Similarity 100.0%; Pred. No. 1.1e-115;  
Matches 241; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 902 GACGACTGGTGTCAACACCACTGTGTGTGACCGTGGTGTGCGGAGACTCCAGAA 961  
Db 402 GACGACTGGTGTCTACACACCACTGTGTGTGACCGTGGTGTGCGGAGACTCCAGAA 343  
QY 962 CTTGGGAGAGGCTGCAAAAAGACCGCCAGAAAGCGGAGAGCTGGCGGTCTCAACCTG 1021  
Db 342 CTTGGGAGAGGCTGCAAAAAGACCGCCAGAAAGCGGAGAGCTGGCGGTCTCAACCTG 283  
QY 1022 CGCCGGCTGACTGTGTGTGCTGCGGACCGGGGCTGGCCCGCGAGAGCGCGGAGTT 1081  
Db 282 CGCCGGCTGACTGTGTGTGCTGCGGACCGGGGCTGGCCCGCGAGAGCGCGGAGTT 223  
QY 1082 CGAGCGGCTGTGGTGGCTTCTCGGGCTGCTGGAACCTGTGAAAGCGGACATGCAAG 1141  
Db 222 CGAGCGGCTGTGGTGGCTTCTCGGGCTGCTGGAACCTGTGAAAGCGGACATGCAAG 163  
QY 1142 C 1142  
Db 162 C 162

## RESULT 9

US-09-908-975-22330  
; Sequence 22330, Application US/09908975  
; Publication No. US20030165843A1  
; GENERAL INFORMATION:  
; APPLICANT: SHOSHAN, Avi  
; APPLICANT: WASSERMAN, Alon  
; APPLICANT: MINTZ, Eli  
; APPLICANT: FAIGLER, Simchon  
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLIC

```

; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME
; FILE REFERENCE: 36688-0005
; CURRENT APPLICATION NUMBER: US/09/908,975
; CURRENT FILING DATE: 2001-07-20
; PRIOR APPLICATION NUMBER: US 60/287,724
; PRIOR FILING DATE: 2001-05-02
; PRIOR APPLICATION NUMBER: US 60/221,607
; PRIOR FILING DATE: 2000-07-28
; NUMBER OF SEQ ID NOS: 32337
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 23330
; LENGTH: 60
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-908-975-23330

Query Match          1.9%; Score 60; DB 3; Length 60;
Best Local Similarity 100.0%; Pred. No. 3.2e-20;
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2464 GGGCTGGGAACTCTCTTAGAGCACTTATCTTATCCCTCGAANGCGTGG 2523
DB      1 GGGCTGGGAACTCTCTTAGAGCACTTATCTTATCCCTCGAANGCGTGG 60

RESULT 10
US-10-723-860-1753/c
; Sequence 1753, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Gineburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05882, 0193, NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1753
; LENGTH: 175737
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-723-860-1753

Query Match          1.7%; Score 53; DB 8; Length 175737;
Best Local Similarity 100.0%; Pred. No. 1.3e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2888 TGAGGCAAGTGATCACTGAGGCCAGAGTTGAGAGCCAGCCTGGCCAAACAT 2940
DB      47829 TGAGGCAAGTGATCACTGAGGCCAGAGTTGAGAGCCAGCCTGGCCAAACAT 47777

RESULT 11
US-10-783-271-34/c
; Sequence 34, Application US/10783271
; Publication No. US20050186577A1
; GENERAL INFORMATION:
; APPLICANT: Veridex, LLC
; APPLICANT: Wang, Yixin
; TITLE OF INVENTION: BREAST CANCER PROGNOSTICS
; FILE REFERENCE: VDX-5003 USP
; CURRENT APPLICATION NUMBER: US/10/783,271
; CURRENT FILING DATE: 2004-02-20
; NUMBER OF SEQ ID NOS: 111
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 34
; LENGTH: 175737
; TYPE: DNA
```

```

; ORGANISM: human
US-10-783-271-34

Query Match          1.7%; Score 53; DB 9; Length 175737;
Best Local Similarity 100.0%; Pred. No. 1.3e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2888 TGAGGCAAGTGATCACTGAGGCCAGAGTTGAGAGCCAGCCTGGCCAAACAT 2940
DB      47829 TGAGGCAAGTGATCACTGAGGCCAGAGTTGAGAGCCAGCCTGGCCAAACAT 47777

RESULT 12
US-10-674-124A-76/c
; Sequence 76, Application US/10674124A
; Publication No. US2004019797A1
; GENERAL INFORMATION:
; APPLICANT: INOKO, Hideooshi
; APPLICANT: TAMURA, Gen
; TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE
; FILE REFERENCE: ORIN-003CIP
; CURRENT APPLICATION NUMBER: US/10/674,124A
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: 10/257,511
; PRIOR FILING DATE: 2003-03-07
; PRIOR APPLICATION NUMBER: PCT/JP00/07621
; PRIOR FILING DATE: 2000-10-30
; PRIOR APPLICATION NUMBER: JP2000-112699
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: JP2002-327516
; PRIOR FILING DATE: 2002-09-28
; PRIOR APPLICATION NUMBER: JP2002-383869
; PRIOR FILING DATE: 2002-12-09
; NUMBER OF SEQ ID NOS: 27110
; SEQ ID NO 76
; LENGTH: 428
; TYPE: DNA
; ORGANISM: Homo sapiens
; PEATRE:
; OTHER INFORMATION: AL357552.12_93957
; FEATURE:
; OTHER INFORMATION: Located on chromosome 1
; FEATURE:
; OTHER INFORMATION: Distance between a terminus base of telomere on
; OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base
; OTHER INFORMATION: sequence : 8481068
; FEATURE:
; OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of
; OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and
; OTHER INFORMATION: 5'-terminus of this base sequence : 60131
US-10-674-124A-76

Query Match          1.7%; Score 52; DB 8; Length 428;
Best Local Similarity 100.0%; Pred. No. 5e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3071 CAAGTTGTGCGACCTGCACTCGAGCTGGGCAACAGCAAGACTCTCTTC 3122
DB      167 CAAGTTGTGCGACCTGCACTCGAGCTGGGCAACAGCAAGACTCTCTTC 116

RESULT 13
US-09-925-065A-591947/c
; Sequence 591947, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
```

```
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 591947
/ LENGTH: 559
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-591947
```

```
Query Match 1.7% Score 52; DB 4; Length 559;
Best Local Similarity 100.0%; Pred. No. 5e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 3071 CAAGATTGTGCCTGACTCCAGCTGGGCAACAGACCAAGCTGTCTC 3122
Db 164 CAAGATTGTGCCTGACTCCAGCTGGGCAACAGACCAAGCTGTCTC 113
```

```
RESULT 14
US-10-027-632-41688
/ Sequence 41688, Application US/10027632
/ Publication No. US20020198371A1
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 41688
/ LENGTH: 568
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-41688
```

```
Query Match 1.7% Score 52; DB 5; Length 568;
Best Local Similarity 100.0%; Pred. No. 5e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2889 GAGGCAAGTGATCACTGAGGCCAGAGATTGAGACCAAGCTGGCCAACT 2940
Db 328 GAGGCAAGTGATCACTGAGGCCAGAGATTGAGACCAAGCTGGCCAACT 379
```

```
RESULT 15
US-10-027-632-41688
/ Sequence 41688, Application US/10027632
/ Publication No. US20030204075A9
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 41688
/ LENGTH: 568
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-41688
```

```
Query Match 1.7% Score 52; DB 6; Length 568;
Best Local Similarity 100.0%; Pred. No. 5e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2889 GAGGCAAGTGATCACTGAGGCCAGAGATTGAGACCAAGCTGGCCAACT 2940
Db 328 GAGGCAAGTGATCACTGAGGCCAGAGATTGAGACCAAGCTGGCCAACT 379
```

```
RESULT 16
US-09-925-065A-768118
/ Sequence 768118, Application US/09925065A
/ Publication No. US20050228172A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 768118
/ LENGTH: 574
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-768118
```

```
Query Match 1.7% Score 52; DB 4; Length 574;
Best Local Similarity 100.0%; Pred. No. 5e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2889 GAGGCAAGTGATCACTGAGGCCAGAGATTGAGACCAAGCTGGCCAACT 2940
```

Db 235 GAGGAGGTGATCACTGAGGCCAGAGATTGAGACAGCCTGGCCAAAT 286

RESULT 17  
US-09-925-065A-59682/c  
; Sequence 59682, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 59682  
; LENGTH: 592  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-59682

Query Match 1.7%; Score 52; DB 4; Length 592;  
Best Local Similarity 100.0%; Pred. No. 5e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGAGGTGATCACTGAGGCCAGAGATTGAGACAGCCTGGCCAAAT 2940  
Db 175 GAGGAGGTGATCACTGAGGCCAGAGATTGAGACAGCCTGGCCAAAT 124

RESULT 18  
US-09-925-065A-386361/c  
; Sequence 386361, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 386361  
; LENGTH: 617  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-386361

Query Match 1.7%; Score 52; DB 4; Length 617;  
Best Local Similarity 100.0%; Pred. No. 5e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTCACCTGCACTCCAGGCTGGGCAACAGAGCAACTGTGCTC 3122  
Db 306 CAAGATTGTCACCTGCACTCCAGGCTGGGCAACAGAGCAACTGTGCTC 255

RESULT 19  
US-10-027-632-61987  
; Sequence 61987, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 61987  
; LENGTH: 631  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-61987

Query Match 1.7%; Score 52; DB 5; Length 631;  
Best Local Similarity 100.0%; Pred. No. 5e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGAGGTGATCACTGAGGCCAGAGATTGAGACAGCCTGGCCAAAT 2940  
Db 361 GAGGAGGTGATCACTGAGGCCAGAGATTGAGACAGCCTGGCCAAAT 412

RESULT 20  
US-10-027-632-61988  
; Sequence 61988, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1999-08-09



```
Publication No. US20030204075A9
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
FILE REFERENCE: 108827.129
CURRENT APPLICATION NUMBER: US/10/027,632
CURRENT FILING DATE: 2002-04-30
PRIOR APPLICATION NUMBER: US 60/218,006
PRIOR FILING DATE: 2000-07-12
PRIOR APPLICATION NUMBER: US 60/198,676
PRIOR FILING DATE: 2000-04-20
PRIOR APPLICATION NUMBER: US 60/193,483
PRIOR FILING DATE: 2000-03-29
PRIOR APPLICATION NUMBER: US 60/185,218
PRIOR FILING DATE: 2000-02-24
PRIOR APPLICATION NUMBER: US 60/167,363
PRIOR FILING DATE: 1999-11-23
PRIOR APPLICATION NUMBER: US 60/156,358
PRIOR FILING DATE: 1999-09-28
PRIOR APPLICATION NUMBER: US 60/146,002
PRIOR FILING DATE: 1999-08-09
NUMBER OF SEQ ID NOS: 325720
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 61989
LENGTH: 631
TYPE: DNA
ORGANISM: Human
US-10-027-632-61989

Query Match
Best Local Similarity 1.7%; Score 52; DB 6; Length 631;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCGATGATCACTGAGGCGGAGGCTTCCAGACCAAGCTGGCCAAAT 2940
DB 361 GAGGCGATGATCACTGAGGCGGAGGCTTCCAGACCAAGCTGGCCAAAT 412

RESULT 25
US-09-925-065A-708701/c
Sequence 708701, Application US/09925065A
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
CURRENT FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 708701
LENGTH: 695
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-708701

Query Match
Best Local Similarity 1.7%; Score 52; DB 4; Length 695;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCGATGATCACTGAGGCGGAGGCTTCCAGACCAAGCTGGCCAAAT 2940
```

```
DB 490 GAGGCGATGATCACTGAGGCGGAGGCTTCCAGACCAAGCTGGCCAAAT 439

RESULT 26
US-10-450-763-14141/c
Sequence 14141, Application US/10450763
Publication No. US20050196754A1
GENERAL INFORMATION:
APPLICANT: HySeq, Inc.
TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
FILE REFERENCE: 790CIP3/US
CURRENT APPLICATION NUMBER: US/10/450,763
CURRENT FILING DATE: 2003-06-11
PRIOR APPLICATION NUMBER: PCT/US01/08631
PRIOR FILING DATE: 2001-03-30
PRIOR APPLICATION NUMBER: 09/540,217
PRIOR FILING DATE: 2000-03-31
PRIOR APPLICATION NUMBER: 09/649,167
PRIOR FILING DATE: 2000-08-23
NUMBER OF SEQ ID NOS: 60736
SOFTWARE: Custom
SEQ ID NO 14141
LENGTH: 1437
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: SIMILAR
LOCATION: (253)..(786)
OTHER INFORMATION: 50% homologous to Homo sapiens capping protein alpha subunit
OTHER INFORMATION: Isoform 1, accession number U56377, Smith-Waterman Score=380.
US-10-450-763-14141

Query Match
Best Local Similarity 1.7%; Score 52; DB 9; Length 1437;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCACTGCACTCCAGCTGGGCAAGAGCAAGACTGTCTC 3122
DB 243 CAAGATTGTGCACTGCACTCCAGCTGGGCAAGAGCAAGACTGTCTC 192

RESULT 27
US-09-764-891-8895/c
Sequence 8895, Application US/09764891
Publication No. US2003007808A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PAM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8895
LENGTH: 9620
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-891-8895

Query Match
Best Local Similarity 1.7%; Score 52; DB 3; Length 9620;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCACTGCACTCCAGCTGGGCAAGAGCAAGACTGTCTC 3122
DB 7058 CAAGATTGTGCACTGCACTCCAGCTGGGCAAGAGCAAGACTGTCTC 7007

RESULT 28
US-09-918-686-1
Sequence 1, Application US/09918686
```



```
; Patent No. US20020076720A1
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary
; APPLICANT: Proll, Sean
; APPLICANT: Paepfer, Bryan
; APPLICANT: Staehling-Hampton, Karen
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; FILE REFERENCE: 240083.515
; CURRENT APPLICATION NUMBER: US/09/918,686
; CURRENT FILING DATE: 2001-07-30
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 92139
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 7043..8369, 8401
; OTHER INFORMATION: n = A,T,C or G
US-09-918-686-1
```

```
Query Match      1.7%; Score 52; DB 3; Length 92139;
Best Local Similarity 100.0%; Pred. No. 4.4e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2889 GAGGCAAGTGTGATCACTGAGGCCAGAGTTGAGACCCAGCTGGCCAACAT 2940
DB      85294 GAGGCAAGTGTGATCACTGAGGCCAGAGTTGAGACCCAGCTGGCCAACAT 85345
```

```
RESULT 29
US-10-353-150-1
; Sequence 1, Application US/10353150
; Publication No. US20030157543A1
; GENERAL INFORMATION:
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Proll, Sean
; APPLICANT: Paepfer, Bryan
; APPLICANT: Staehling-Hampton, Karen
; TITLE OF INVENTION: METHODS FOR IDENTIFYING
; FILE REFERENCE: 240083.515C1
; CURRENT APPLICATION NUMBER: US/10/353,150
; CURRENT FILING DATE: 2003-01-27
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 92139
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 7043..8369, 8401
; OTHER INFORMATION: n = A,T,C or G
US-10-353-150-1
```

```
Query Match      1.7%; Score 52; DB 6; Length 92139;
Best Local Similarity 100.0%; Pred. No. 4.4e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2889 GAGGCAAGTGTGATCACTGAGGCCAGAGTTGAGACCCAGCTGGCCAACAT 2940
DB      85294 GAGGCAAGTGTGATCACTGAGGCCAGAGTTGAGACCCAGCTGGCCAACAT 85345
```

```
RESULT 30
US-10-408-168-1/C
; Sequence 1, Application US/10408168
; Publication No. US20030235847A1
; GENERAL INFORMATION:
; APPLICANT: Paepfer, Bryan W.
```

```
; APPLICANT: Proll, Sean
; APPLICANT: Charnley, Patrick R.
; APPLICANT: Brunkow, Mary E.
; APPLICANT: Utterlinden, Andreas Gerardus
; TITLE OF INVENTION: ASSOCIATION OF POLYMORPHISMS IN THE SOST
; FILE REFERENCE: 240083.525
; CURRENT APPLICATION NUMBER: US/10/408,168
; CURRENT FILING DATE: 2003-04-03
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 130320
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 100722, 100754, 102080, 117731, 124408, 124532, 124585, 124955,
; LOCATION: 124956, 124963, 124964, 124965, 124967
; OTHER INFORMATION: n = A,T,C or G
US-10-408-168-1
```

```
Query Match      1.7%; Score 52; DB 6; Length 130320;
Best Local Similarity 100.0%; Pred. No. 4.3e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2889 GAGGCAAGTGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAACAT 2940
DB      23819 GAGGCAAGTGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAACAT 23768
```

```
RESULT 31
US-10-085-117-244
; Sequence 244, Application US/10085117
; Publication No. US2003023234A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 244
; LENGTH: 136726
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: variation
; LOCATION: (1)...(136726)
; OTHER INFORMATION: n = any nucleotide
US-10-085-117-244
```

```
Query Match      1.7%; Score 52; DB 6; Length 136726;
Best Local Similarity 100.0%; Pred. No. 4.3e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2889 GAGGCAAGTGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAACAT 2940
DB      66522 GAGGCAAGTGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAACAT 66573
```

```
RESULT 32
US-10-981-277-50
; Sequence 50, Application US/10981277
; Publication No. US20050181389A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; APPLICANT: Davis, Lisa
; TITLE OF INVENTION: Compositions and Methods for Glioma Classification
```

```
/ FILE REFERENCE: 03-968-US
/ CURRENT APPLICATION NUMBER: US/10/981,277
/ CURRENT FILING DATE: 2004-11-03
/ PRIOR APPLICATION NUMBER: US 60/516,817
/ PRIOR FILING DATE: 2003-11-03
/ NUMBER OF SEQ ID NOS: 57
/ SOFTWARE: PatentIn version 3.3
/ SEQ ID NO 50
/ LENGTH: 150437
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-10-981-277-50

Query Match      1.7%: Score 52; DB 9; Length 150437;
Best Local Similarity 100.0%: Pred. No. 4,3e-16;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 3122
Db      125751 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTGTCTC 125802

RESULT 33
/ US-10-174-014-12
/ Sequence 12, Application US/10174014
/ Publication No. US2004005292A1
/ GENERAL INFORMATION:
/ APPLICANT: C. Frank Bennett
/ APPLICANT: Susan M. Freiler
/ APPLICANT: Kenneth W. Dobie
/ TITLE OF INVENTION: ANTISENSE MODULATION OF SMRT EXPRESSION
/ FILE REFERENCE: PTS-0012
/ CURRENT APPLICATION NUMBER: US/10/174,014
/ CURRENT FILING DATE: 2002-06-17
/ NUMBER OF SEQ ID NOS: 73
/ SEQ ID NO 12
/ LENGTH: 221000
/ TYPE: DNA
/ ORGANISM: H. sapiens
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77967
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77968
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77969
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77970
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77971
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77972
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77973
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77974
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
```

```
/ LOCATION: 77975
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77976
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77977
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77978
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77979
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77980
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77981
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77982
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77983
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77984
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77985
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77986
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77987
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77988
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77989
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77990
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77991
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77992
/ OTHER INFORMATION: unknown
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: 77993
```

OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 77994  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 77995  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 77996  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 77997  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 77998  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 77999  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78000  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78001  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78002  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78003  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78004  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78005  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78006  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78007  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78008  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78009  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78010  
OTHER INFORMATION: unknown  
FEATURE:  
NAME/KEY: unsure  
LOCATION: 78011  
OTHER INFORMATION: unknown

FEATURE:  
NAME/KEY: unsure  
LOCATION: 78012  
OTHER INFORMATION: unknown  
FEATURE:

Query Match 1.7% Score 52; DB 6; Length 221000;  
Best Local Similarity 100.0%; Pred. No. 4.3e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCACCTGCACCTCCAGCTGGGCAACAGCAAGACTGTGTCTC 3122  
Db 173223 CAAGATTGTGCACCTGCACCTCCAGCTGGGCAACAGCAAGACTGTGTCTC 173274

RESULT 34  
US-10-087-192-652  
Sequence 652, Application US/10087192  
Publication No. US20020182586A1  
GENERAL INFORMATION:  
APPLICANT: Morris, David W.  
APPLICANT: Engelhard, Eric K.  
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
FILE REFERENCE: 529452000122  
CURRENT APPLICATION NUMBER: US/10/087,192  
CURRENT FILING DATE: 2002-03-01  
PRIOR APPLICATION NUMBER: US 09/747,377  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: US 09/798,586  
PRIOR FILING DATE: 2001-03-02  
NUMBER OF SEQ ID NOS: 2059  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 652  
LENGTH: 233380  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(233380)  
OTHER INFORMATION: n = A,T,C or G  
US-10-087-192-652

Query Match 1.7% Score 52; DB 5; Length 233380;  
Best Local Similarity 100.0%; Pred. No. 4.3e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCACCTGCACCTCCAGCTGGGCAACAGCAAGACTGTGTCTC 3122  
Db 179881 CAAGATTGTGCACCTGCACCTCCAGCTGGGCAACAGCAAGACTGTGTCTC 179932

RESULT 35  
US-10-741-601-5719/c  
Sequence 5719, Application US/10741601  
Publication No. US2004016519A1  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
FILE REFERENCE: CL001500  
CURRENT APPLICATION NUMBER: US/10/741,601  
CURRENT FILING DATE: 2003-12-22  
NUMBER OF SEQ ID NOS: 26415  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 5719  
LENGTH: 398287  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(398287)  
OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1

US-10-741-601-5719

Query Match 1.7%; Score 52; DB 7; Length 398287;  
Best Local Similarity 100.0%; Pred.No. 4.2e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGGCACTGCACTCCAGCTGGGCAACAGAGAACTGTCTC 3122  
DB 281811 CAAGATTGGCACTGCACTCCAGCTGGGCAACAGAGAACTGTCTC 281760

RESULT 36  
US-10-741-600-17839/c  
; Sequence 17839, Application US/10741600  
; Publication No. US20050026169A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CL001499  
; CURRENT APPLICATION NUMBER: US/10/741, 600  
; PRIOR FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 73997  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 17839  
; LENGTH: 398287  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...((398287))  
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-  
US-10-741-600-17839

Query Match 1.7%; Score 52; DB 8; Length 398287;  
Best Local Similarity 100.0%; Pred.No. 4.2e-16;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGGCACTGCACTCCAGCTGGGCAACAGAGAACTGTCTC 3122  
DB 281811 CAAGATTGGCACTGCACTCCAGCTGGGCAACAGAGAACTGTCTC 281760

RESULT 37  
US-09-925-065A-211249/c  
; Sequence 211249, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925, 065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243, 096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252, 147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250, 092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261, 766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289, 846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 211249  
; LENGTH: 385  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-211249

Query Match 1.6%; Score 51; DB 4; Length 385;

Best Local Similarity 100.0%; Pred.No. 1.7e-15;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2890 AGGCAGGTGATCACTGAGGCGCAGAGTTTCAGACCAAGCTGGCCAAAT 2940  
DB 167 AGGCAGGTGATCACTGAGGCGCAGAGTTTCAGACCAAGCTGGCCAAAT 117

RESULT 38  
US-09-925-065A-141833  
; Sequence 141833, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925, 065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243, 096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252, 147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250, 092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261, 766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289, 846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 141833  
; LENGTH: 561  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-141833

Query Match 1.6%; Score 51; DB 4; Length 561;  
Best Local Similarity 100.0%; Pred.No. 1.7e-15;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2890 AGGCAGGTGATCACTGAGGCGCAGAGTTTCAGACCAAGCTGGCCAAAT 2940  
DB 402 AGGCAGGTGATCACTGAGGCGCAGAGTTTCAGACCAAGCTGGCCAAAT 452

RESULT 39  
US-10-027-632-54553  
; Sequence 54553, Application US/10027632  
; Publication No. US20020198371A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027, 632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218, 006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198, 676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193, 483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185, 218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167, 363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156, 358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146, 002  
; PRIOR FILING DATE: 1999-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 54553  
LENGTH: 622  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-54553

Query Match 1.6%; Score 51; DB 5; Length 622;  
Best Local Similarity 100.0%; Pred. No. 1.7e-15;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2890 AGGACGTGATCCTGAGGCCAGAGTTCCAGACCACTGGCCCAACAT 2940  
DB 401 AGGACGTGATCCTGAGGCCAGAGTTCCAGACCACTGGCCCAACAT 451

RESULT 40  
US-10-027-632-55034  
Sequence 55034, Application US/10027632  
Publication No. US20020198371A1  
GENERAL INFORMATION:

APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
POLYMORPHISMS IN THE HUMAN GENOME  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
PRIOR FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 55034  
LENGTH: 622  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-55034

Query Match 1.6%; Score 51; DB 5; Length 622;  
Best Local Similarity 100.0%; Pred. No. 1.7e-15;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2890 AGGACGTGATCCTGAGGCCAGAGTTCCAGACCACTGGCCCAACAT 2940  
DB 401 AGGACGTGATCCTGAGGCCAGAGTTCCAGACCACTGGCCCAACAT 451

RESULT 41  
US-10-027-632-56282  
Sequence 56282, Application US/10027632  
Publication No. US20020198371A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
POLYMORPHISMS IN THE HUMAN GENOME  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
PRIOR FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 56282  
LENGTH: 622  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-56282

Query Match 1.6%; Score 51; DB 5; Length 622;  
Best Local Similarity 100.0%; Pred. No. 1.7e-15;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2890 AGGACGTGATCCTGAGGCCAGAGTTCCAGACCACTGGCCCAACAT 2940  
DB 401 AGGACGTGATCCTGAGGCCAGAGTTCCAGACCACTGGCCCAACAT 451

RESULT 42  
US-10-027-632-56384  
Sequence 56384, Application US/10027632  
Publication No. US20020198371A1  
GENERAL INFORMATION:

APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
POLYMORPHISMS IN THE HUMAN GENOME  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
PRIOR FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 56384  
LENGTH: 622  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-56384

Query Match 1.6%; Score 51; DB 5; Length 622;  
Best Local Similarity 100.0%; Pred. No. 1.7e-15;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2890 AGGACGTGATCCTGAGGCCAGAGTTCCAGACCACTGGCCCAACAT 2940  
DB 401 AGGACGTGATCCTGAGGCCAGAGTTCCAGACCACTGGCCCAACAT 451

RESULT 43  
US-10-027-632-180536  
Sequence 180536, Application US/10027632  
Publication No. US20020198371A1  
GENERAL INFORMATION:

```

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 180536
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-180536
```

```
Query Match      1.6%; Score 51; DB 5; Length 622;
Best Local Similarity 100.0%; Pred. No. 1.7e-15;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      2890 AGGCGGTGATCACCTGAGGCGCAGAGTTGAGACCAAGCTGGCCAAAT 2940
Db      401 AGGCGGTGATCACCTGAGGCGCAGAGTTGAGACCAAGCTGGCCAAAT 451
```

```

RESULT 44
US-10-027-632-314485
; Sequence 314485, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 314485
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-314485
```

```
Query Match      1.6%; Score 51; DB 5; Length 622;
Best Local Similarity 100.0%; Pred. No. 1.7e-15;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      2890 AGGCGGTGATCACCTGAGGCGCAGAGTTGAGACCAAGCTGGCCAAAT 2940
Db      401 AGGCGGTGATCACCTGAGGCGCAGAGTTGAGACCAAGCTGGCCAAAT 451
```

```

RESULT 45
US-10-027-632-54553
; Sequence 54553, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 54553
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-54553
```

```
Query Match      1.6%; Score 51; DB 5; Length 622;
Best Local Similarity 100.0%; Pred. No. 1.7e-15;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      2890 AGGCGGTGATCACCTGAGGCGCAGAGTTGAGACCAAGCTGGCCAAAT 2940
Db      401 AGGCGGTGATCACCTGAGGCGCAGAGTTGAGACCAAGCTGGCCAAAT 451
```

```

RESULT 46
US-10-027-632-55034
; Sequence 55034, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT FILING DATE: 2002-04-30
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-07-12
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-04-20
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-03-29
; PRIOR FILING DATE: 2000-02-24
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-11-23
; PRIOR FILING DATE: 1999-09-28
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 55034
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-55034
```

```

; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 55034
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-55034

Query Match          1.6%; Score 51; DB 6; Length 622;
Best Local Similarity 100.0%; Pred. No. 1.7e-15;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      2890 AGGCAAGTGTGATCACCCTGAGGCGCAGAGTTGCAGACCAAGCTGGCCAAACAT 2940
DB      401 AGGCAAGTGTGATCACCCTGAGGCGCAGAGTTGCAGACCAAGCTGGCCAAACAT 451

RESULT 47
US-10-027-632-56282
; Sequence 56282, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56282
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-56282

Query Match          1.6%; Score 51; DB 6; Length 622;
Best Local Similarity 100.0%; Pred. No. 1.7e-15;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      2890 AGGCAAGTGTGATCACCCTGAGGCGCAGAGTTGCAGACCAAGCTGGCCAAACAT 2940
DB      401 AGGCAAGTGTGATCACCCTGAGGCGCAGAGTTGCAGACCAAGCTGGCCAAACAT 451

RESULT 48
US-10-027-632-56384
; Sequence 56384, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
```

```

; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 56384
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-56384

Query Match          1.6%; Score 51; DB 6; Length 622;
Best Local Similarity 100.0%; Pred. No. 1.7e-15;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      2890 AGGCAAGTGTGATCACCCTGAGGCGCAGAGTTGCAGACCAAGCTGGCCAAACAT 2940
DB      401 AGGCAAGTGTGATCACCCTGAGGCGCAGAGTTGCAGACCAAGCTGGCCAAACAT 451

RESULT 49
US-10-027-632-180536
; Sequence 180536, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 180536
; LENGTH: 622
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-180536

Query Match          1.6%; Score 51; DB 6; Length 622;
Best Local Similarity 100.0%; Pred. No. 1.7e-15;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      2890 AGGCAAGTGTGATCACCCTGAGGCGCAGAGTTGCAGACCAAGCTGGCCAAACAT 2940
DB      401 AGGCAAGTGTGATCACCCTGAGGCGCAGAGTTGCAGACCAAGCTGGCCAAACAT 451

RESULT 50
US-10-027-632-314485
; Sequence 314485, Application US/10027632
```



```
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 314485
/ LENGTH: 622
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-314485
```

```
Query Match      1.6%; Score 51; DB 6; Length 622;
Best Local Similarity 100.0%; Pred. No. 1.7e-15;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      2890 AGGCAAGTGATGATCAGCTGAGGCGCAGAGTTTCAGACCACTGGGCCCAACAT 2940
Db      401 AGGCAAGTGATGATCAGCTGAGGCGCAGAGTTTCAGACCACTGGGCCCAACAT 451
```

```
RESULT 51
US-09-925-298-220/c
/ Sequence 220, Application US/09925298
/ Publication No. US20020039764A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
/ FILE REFERENCE: PA103
/ CURRENT APPLICATION NUMBER: US/09/925,298
/ CURRENT FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: PCT/US00/05881
/ PRIOR FILING DATE: 2000-03-08
/ PRIOR APPLICATION NUMBER: 60/124,270
/ PRIOR FILING DATE: 1999-03-12
/ NUMBER OF SEQ ID NOS: 846
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 220
/ LENGTH: 832
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-298-220
```

```
Query Match      1.6%; Score 51; DB 3; Length 832;
Best Local Similarity 100.0%; Pred. No. 1.7e-15;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      2890 AGGCAAGTGATGATCAGCTGAGGCGCAGAGTTTCAGACCACTGGGCCCAACAT 2940
Db      76 AGGCAAGTGATGATCAGCTGAGGCGCAGAGTTTCAGACCACTGGGCCCAACAT 26
```

```
RESULT 52
US-10-102-806-220/c
/ Sequence 220, Application US/10102806
```

```
/ Publication No. US20030054421A1
/ GENERAL INFORMATION:
/ APPLICANT: Rosen et al.
/ TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
/ FILE REFERENCE: PA103P1C1
/ CURRENT APPLICATION NUMBER: US/10/102,806
/ CURRENT FILING DATE: 2002-03-22
/ PRIOR APPLICATION NUMBER: 09/925,298
/ PRIOR FILING DATE: 2001-08-10
/ PRIOR APPLICATION NUMBER: PCT/US00/05881
/ PRIOR FILING DATE: 2000-03-08
/ PRIOR APPLICATION NUMBER: 60/124,270
/ PRIOR FILING DATE: 1999-03-12
/ NUMBER OF SEQ ID NOS: 846
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 220
/ LENGTH: 832
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-102-806-220
```

```
Query Match      1.6%; Score 51; DB 5; Length 832;
Best Local Similarity 100.0%; Pred. No. 1.7e-15;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      2890 AGGCAAGTGATGATCAGCTGAGGCGCAGAGTTTCAGACCACTGGGCCCAACAT 2940
Db      76 AGGCAAGTGATGATCAGCTGAGGCGCAGAGTTTCAGACCACTGGGCCCAACAT 26
```

```
RESULT 53
US-10-450-763-29534
/ Sequence 29534, Application US/10450763
/ Publication No. US20050196754A1
/ GENERAL INFORMATION:
/ APPLICANT: HySeq, Inc
/ TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
/ FILE REFERENCE: 790CIP3/US
/ CURRENT APPLICATION NUMBER: US/10/450,763
/ CURRENT FILING DATE: 2003-06-11
/ PRIOR APPLICATION NUMBER: PCT/US01/08631
/ PRIOR FILING DATE: 2001-03-30
/ PRIOR APPLICATION NUMBER: 09/540,217
/ PRIOR FILING DATE: 2000-03-31
/ PRIOR APPLICATION NUMBER: 09/649,167
/ PRIOR FILING DATE: 2000-08-23
/ NUMBER OF SEQ ID NOS: 60736
/ SOFTWARE: Custom
/ SEQ ID NO 29534
/ LENGTH: 2791
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ NAME/KEY: SIMILAR
/ LOCATION: (18)...(194)
/ OTHER INFORMATION: 43% homologous to Trypanoplasma borreli ribosomal protein
/ OTHER INFORMATION: S12, accession number U14183, Smith-Waterman Score=95.
US-10-450-763-29534
```

```
Query Match      1.6%; Score 51; DB 9; Length 2791;
Best Local Similarity 100.0%; Pred. No. 1.6e-15;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy      465 CTTAAGGGGAGAGCGCGGCTTTCAGAGCTGTTGCCCGGAGTTGGCACC 515
Db      1538 CTTAAGGGGAGAGCGCGGCTTTCAGAGCTGTTGCCCGGAGTTGGCACC 1588
```

```
RESULT 54
US-10-737-082-22/c
/ Sequence 22, Application US/10737082
/ Publication No. US20050130170A1
/ GENERAL INFORMATION:
```

APPLICANT: Bayer Healthcare LLC  
APPLICANT: Beard, Chris  
APPLICANT: Burgess, Chris  
APPLICANT: Gannon, Allison  
APPLICANT: Harvey, Jeanne  
APPLICANT: Lechner, John F.  
TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences  
FILE REFERENCE: 1657/2032  
CURRENT FILING DATE: US/10/737,082  
PRIOR FILING DATE: 2003-12-16  
PRIOR APPLICATION NUMBER: US 10/737,082  
NUMBER OF SEQ ID NOS: 300  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 22  
LENGTH: 23139  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-737-082-22

Query Match 1.6%; Score 51; DB 9; Length 23139;  
Best Local Similarity 100.0%; Pred. No. 1.5e-15;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2890 AGGCAGGTGATCACCCTGAGCCAGAGTTCGAGACCAAGCTGGCCACAT 2940  
DB 3783 AGGCAGGTGATCACCCTGAGCCAGAGTTCGAGACCAAGCTGGCCACAT 3733

RESULT 55  
US-10-765-790-22/c  
Sequence 22, Application US/10765790  
GENERAL INFORMATION:  
APPLICANT: Bayer Healthcare LLC  
APPLICANT: Beard, Chris  
APPLICANT: Burgess, Chris  
APPLICANT: Gannon, Allison  
APPLICANT: Harvey, Jeanne  
APPLICANT: Lechner, John F.  
TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences  
FILE REFERENCE: 1657/2035  
CURRENT FILING DATE: US/10/765,790  
CURRENT FILING DATE: 2004-01-27  
PRIOR FILING DATE: 2003-12-16  
PRIOR APPLICATION NUMBER: US 10/737,082  
NUMBER OF SEQ ID NOS: 300  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 22  
LENGTH: 23139  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-765-790-22

Query Match 1.6%; Score 51; DB 9; Length 23139;  
Best Local Similarity 100.0%; Pred. No. 1.5e-15;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2890 AGGCAGGTGATCACCCTGAGCCAGAGTTCGAGACCAAGCTGGCCACAT 2940  
DB 3783 AGGCAGGTGATCACCCTGAGCCAGAGTTCGAGACCAAGCTGGCCACAT 3733

RESULT 56  
US-10-840-590-3/c  
Sequence 3, Application US/10840590  
GENERAL INFORMATION:  
APPLICANT: ROTH, RICHARD B.  
APPLICANT: NELSON, MATTHEW ROBERTS  
APPLICANT: KAMMERER, STEFAN M.

APPLICANT: BRAUN, ANDREAS  
APPLICANT: HOYAL-WRIGHTSON, CAROLYN R.  
TITLE OF INVENTION: METHODS FOR IDENTIFYING RISK OF MELANOMA AND TREATMENTS THEREOF  
FILE REFERENCE: SEQ-4061-CP  
CURRENT FILING DATE: US/10/840,590  
CURRENT FILING DATE: 2004-05-05  
PRIOR FILING DATE: 2003-07-23  
PRIOR APPLICATION NUMBER: 60/489,703  
NUMBER OF SEQ ID NOS: 1638  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 3  
LENGTH: 68200  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (50036)..(50036)  
OTHER INFORMATION: n 1s a, c, g, or t  
US-10-840-590-3

Query Match 1.6%; Score 51; DB 9; Length 68200;  
Best Local Similarity 100.0%; Pred. No. 1.5e-15;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2890 AGGCAGGTGATCACCCTGAGCCAGAGTTCGAGACCAAGCTGGCCACAT 2940  
DB 63277 AGGCAGGTGATCACCCTGAGCCAGAGTTCGAGACCAAGCTGGCCACAT 63227

RESULT 57  
US-10-087-192-1360/c  
Sequence 1360, Application US/10087192  
GENERAL INFORMATION:  
APPLICANT: Morris, David W.  
APPLICANT: Engelhard, Eric K.  
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
FILE REFERENCE: 529452000122  
CURRENT FILING DATE: US/10/087,192  
CURRENT FILING DATE: 2002-03-01  
PRIOR FILING DATE: 2000-12-22  
PRIOR APPLICATION NUMBER: US 09/747,377  
PRIOR FILING DATE: 2001-03-02  
NUMBER OF SEQ ID NOS: 2059  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1360  
LENGTH: 215221  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-087-192-1360

Query Match 1.6%; Score 51; DB 5; Length 215221;  
Best Local Similarity 100.0%; Pred. No. 1.4e-15;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGTTGTGCACTGCACTGCAAGCTTGGCAAGAGCAAGACTCTCTCT 3121  
DB 50858 CAAGTTGTGCACTGCACTGCAAGCTTGGCAAGAGCAAGACTCTCTCT 50808

RESULT 58  
US-09-973-278-870  
Sequence 870, Application US/09973278  
GENERAL INFORMATION:  
APPLICANT: Fischer et al.  
TITLE OF INVENTION: 123 Human Secreted Proteins  
CURRENT FILING DATE: US/09/973,278  
CURRENT FILING DATE: 2001-10-10  
PRIOR APPLICATION NUMBER: 60/239,899

```

PRIOR FILING DATE: 2000-10-13
PRIOR APPLICATION NUMBER: 09/227,357
PRIOR FILING DATE: 1999-01-08
PRIOR APPLICATION NUMBER: PCT/US98/13684
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/051,926
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/052,793
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/051,925
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/051,929
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/052,803
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/052,732
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/051,931
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/051,932
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/051,916
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/051,930
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/051,918
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/051,920
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/052,733
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/052,795
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/051,919
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/051,928
PRIOR FILING DATE: 1997-07-08
PRIOR APPLICATION NUMBER: 60/055,722
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/055,723
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/055,948
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/055,949
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/055,953
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/055,950
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/055,947
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/055,964
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/056,360
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/055,684
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/055,984
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/055,954
PRIOR FILING DATE: 1997-08-18
PRIOR APPLICATION NUMBER: 60/058,785
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,664
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,660
PRIOR FILING DATE: 1997-09-12
PRIOR APPLICATION NUMBER: 60/058,661
PRIOR FILING DATE: 1997-09-12
NUMBER OF SEQ ID NOS: 947
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 870
LENGTH: 288
```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-973-278-870

Query Match
Best Local Similarity 1.6%; Score 50; DB 3; Length 288;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCCACTGCACTTCACGCTGGGCAACAGACGAAGACTGTCTTC 3122
DB 221 AGATTGTGCCACTGCACTTCACGCTGGGCAACAGACGAAGACTGTCTTC 270

RESULT 59
US-10-450-763-29529
; Sequence 29529, Application US/10450763
; Publication No. US20050196754A1
; GENERAL INFORMATION:
; APPLICANT: HySeq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 790CIP3/US
; CURRENT APPLICATION NUMBER: US/10/450,763
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: PCT/US01/08631
; PRIOR FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 29529
; LENGTH: 432
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (413)..(120)
; OTHER INFORMATION: 33% homologous to Human herpesvirus 6 H87, accession number
US-10-450-763-29529

Query Match
Best Local Similarity 1.6%; Score 50; DB 9; Length 432;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 888 GACGGGCTCAACAGACGACTGCGTCTACCAACCACTGTGCTGACCGT 937
DB 22 GACGGGCTCAACAGACGACTGCGTCTACCAACCACTGTGCTGACCGT 71

RESULT 60
US-09-925-065A-484883/c
; Sequence 484883, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 484883
; LENGTH: 571
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-484883

Query Match          1.6%; Score 50; DB 4; Length 571;
Best Local Similarity 100.0%; Pred. No. 5.6e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTGCACCTGCATCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122
DB      271 AGATTGTGCACCTGCATCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 222

RESULT 61
US-09-925-065A-785967/c
; Sequence 785967, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 785967
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-785967

Query Match          1.6%; Score 50; DB 4; Length 572;
Best Local Similarity 100.0%; Pred. No. 5.6e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTGCACCTGCATCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122
DB      347 AGATTGTGCACCTGCATCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 298

RESULT 62
US-09-925-065A-894777/c
; Sequence 894777, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
```

```

; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 894777
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-894777

Query Match          1.6%; Score 50; DB 4; Length 625;
Best Local Similarity 100.0%; Pred. No. 5.6e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTGCACCTGCATCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122
DB      229 AGATTGTGCACCTGCATCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 180

RESULT 63
US-09-925-065A-892664/c
; Sequence 892664, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 892664
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-892664

Query Match          1.6%; Score 50; DB 4; Length 630;
Best Local Similarity 100.0%; Pred. No. 5.6e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTGCACCTGCATCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122
DB      234 AGATTGTGCACCTGCATCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 185

RESULT 64
US-09-925-065A-892405/c
; Sequence 892405, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
```

```
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 892405
/ LENGTH: 633
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-892405

Query Match          1.6%; Score 50; DB 4; Length 633;
Best Local Similarity 100.0%; Pred. No. 5.6e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3073 AGATTGTGCACCTGCATCTCCAGCTGGGCAACAGACGAACTGTCTC 3122
Db      237 AGATTGTGCACCTGCATCTCCAGCTGGGCAACAGACGAACTGTCTC 188

RESULT 65
US-09-925-065A-892514/c
/ Sequence 892514, Application US/09925065A
/ Publication No. US20050228172A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 892514
/ LENGTH: 633
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-892514

Query Match          1.6%; Score 50; DB 4; Length 633;
Best Local Similarity 100.0%; Pred. No. 5.6e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3073 AGATTGTGCACCTGCATCTCCAGCTGGGCAACAGACGAACTGTCTC 3122
Db      237 AGATTGTGCACCTGCATCTCCAGCTGGGCAACAGACGAACTGTCTC 188

RESULT 66
US-09-925-065A-949927
/ Sequence 949927, Application US/09925065A
/ Publication No. US20050228172A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
```

```
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 949927
/ LENGTH: 633
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-949927

Query Match          1.6%; Score 50; DB 4; Length 633;
Best Local Similarity 100.0%; Pred. No. 5.6e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3073 AGATTGTGCACCTGCATCTCCAGCTGGGCAACAGACGAACTGTCTC 3122
Db      29 AGATTGTGCACCTGCATCTCCAGCTGGGCAACAGACGAACTGTCTC 78

RESULT 67
US-09-925-065A-917225
/ Sequence 917225, Application US/09925065A
/ Publication No. US20050228172A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 917225
/ LENGTH: 676
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-917225

Query Match          1.6%; Score 50; DB 4; Length 676;
Best Local Similarity 100.0%; Pred. No. 5.6e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3073 AGATTGTGCACCTGCATCTCCAGCTGGGCAACAGACGAACTGTCTC 3122
Db      448 AGATTGTGCACCTGCATCTCCAGCTGGGCAACAGACGAACTGTCTC 497

RESULT 68
US-09-925-065A-917226
/ Sequence 917226, Application US/09925065A
/ Publication No. US20050228172A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
```

```

; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 917226
; LENGTH: 676
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-917226

Query Match      1.6%; Score 50; DB 4; Length 676;
Best Local Similarity 100.0%; Pred. No. 5.6e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3073 AGATTGTGCACCTGCACCTCCAGCCTGGGCAAGACAGACAGACTGTGTCTC 3122
DB      448 AGATTGTGCACCTGCACCTCCAGCCTGGGCAAGACAGACAGACTGTGTCTC 497

RESULT 69
US-09-925-065A-944569
; Sequence 944569, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; Nucleotide Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 944569
; LENGTH: 676
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-944569

Query Match      1.6%; Score 50; DB 4; Length 676;
Best Local Similarity 100.0%; Pred. No. 5.6e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3073 AGATTGTGCACCTGCACCTCCAGCCTGGGCAAGACAGACAGACTGTGTCTC 3122
DB      448 AGATTGTGCACCTGCACCTCCAGCCTGGGCAAGACAGACAGACTGTGTCTC 497

RESULT 70
US-10-027-632-158850/c
; Sequence 158850, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```

```

; TITLE OF INVENTION: Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 158850
; LENGTH: 714
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-158850

Query Match      1.6%; Score 50; DB 5; Length 714;
Best Local Similarity 100.0%; Pred. No. 5.6e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3073 AGATTGTGCACCTGCACCTCCAGCCTGGGCAAGACAGACAGACTGTGTCTC 3122
DB      133 AGATTGTGCACCTGCACCTCCAGCCTGGGCAAGACAGACAGACTGTGTCTC 84

RESULT 71
US-10-027-632-158850/c
; Sequence 158850, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; Nucleotide Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 158850
; LENGTH: 714
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-158850

Query Match      1.6%; Score 50; DB 6; Length 714;
Best Local Similarity 100.0%; Pred. No. 5.6e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3073 AGATTGTGCACCTGCACCTCCAGCCTGGGCAAGACAGACAGACTGTGTCTC 3122
```

Db 123 AGATTGTGCACCTGCACCTCCAGCTGTGGCAACAGACAGACTCTGTCTC 84

RESULT 72  
US-09-925-065A-934608

; Sequence 934608, Application US/09925065A  
; Publication No. US20050228172A9

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single

; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.135

; CURRENT APPLICATION NUMBER: US/09/925,065A

; PRIOR FILING DATE: 2001-08-08

; PRIOR APPLICATION NUMBER: US 60/243,096

; PRIOR FILING DATE: 2000-10-24

; PRIOR APPLICATION NUMBER: US 60/252,147

; PRIOR FILING DATE: 2000-11-20

; PRIOR APPLICATION NUMBER: US 60/250,092

; PRIOR FILING DATE: 2000-11-30

; PRIOR APPLICATION NUMBER: US 60/261,766

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/289,846

; PRIOR FILING DATE: 2001-05-09

; NUMBER OF SEQ ID NOS: 957086

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 934608

; LENGTH: 739

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-925-065A-934608

Query Match 1.6%; Score 50; DB 4; Length 739;

Best Local Similarity 100.0%; Pred. No. 5.6e-15;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCCAGCTGTGGCAACAGACAGACTCTGTCTC 3122

Db 603 AGATTGTGCACCTGCACCTCCAGCTGTGGCAACAGACAGACTCTGTCTC 652

RESULT 73  
US-09-925-065A-941367

; Sequence 941367, Application US/09925065A

; Publication No. US20050228172A9

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single

; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.135

; CURRENT APPLICATION NUMBER: US/09/925,065A

; PRIOR FILING DATE: 2001-08-08

; PRIOR APPLICATION NUMBER: US 60/243,096

; PRIOR FILING DATE: 2000-10-24

; PRIOR APPLICATION NUMBER: US 60/252,147

; PRIOR FILING DATE: 2000-11-20

; PRIOR APPLICATION NUMBER: US 60/250,092

; PRIOR FILING DATE: 2000-11-30

; PRIOR APPLICATION NUMBER: US 60/261,766

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/289,846

; PRIOR FILING DATE: 2001-05-09

; NUMBER OF SEQ ID NOS: 957086

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 941367

; LENGTH: 742

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-925-065A-941367

Query Match 1.6%; Score 50; DB 4; Length 742;

Best Local Similarity 100.0%; Pred. No. 5.6e-15;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCCAGCTGTGGCAACAGACAGACTCTGTCTC 3122

Db 603 AGATTGTGCACCTGCACCTCCAGCTGTGGCAACAGACAGACTCTGTCTC 652

RESULT 74  
US-10-027-632-265269

; Sequence 265269, Application US/10027632

; Publication No. US20020198371A1

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; TITLE OF INVENTION: Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; PRIOR FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002

; PRIOR FILING DATE: 1999-08-09

; NUMBER OF SEQ ID NOS: 325720

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 265269

; LENGTH: 1986

; TYPE: DNA

; ORGANISM: Human

US-10-027-632-265269

Query Match 1.6%; Score 50; DB 5; Length 1986;

Best Local Similarity 100.0%; Pred. No. 5.5e-15;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCCAGCTGTGGCAACAGACAGACTCTGTCTC 3122

Db 197 AGATTGTGCACCTGCACCTCCAGCTGTGGCAACAGACAGACTCTGTCTC 246

RESULT 75  
US-10-027-632-265269

; Sequence 265269, Application US/10027632

; Publication No. US20030204075A9

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

; TITLE OF INVENTION: Polymorphisms in the Human Genome

; FILE REFERENCE: 108827.129

; CURRENT APPLICATION NUMBER: US/10/027,632

; PRIOR FILING DATE: 2002-04-30

; PRIOR APPLICATION NUMBER: US 60/218,006

; PRIOR FILING DATE: 2000-07-12

; PRIOR APPLICATION NUMBER: US 60/198,676

; PRIOR FILING DATE: 2000-04-20

; PRIOR APPLICATION NUMBER: US 60/193,483

; PRIOR FILING DATE: 2000-03-29

; PRIOR APPLICATION NUMBER: US 60/185,218

; PRIOR FILING DATE: 2000-02-24

; PRIOR APPLICATION NUMBER: US 60/167,363

; PRIOR FILING DATE: 1999-11-23

; PRIOR APPLICATION NUMBER: US 60/156,358

; PRIOR FILING DATE: 1999-09-28

; PRIOR APPLICATION NUMBER: US 60/146,002



```
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 265269
; LENGTH: 1986
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-265269

Query Match      1.6%; Score 50; DB 6; Length 1986;
Best Local Similarity 100.0%; Pred. No. 5.3e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122
DB      197 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 246

RESULT 76
US-10-374-979-10
; Sequence 10, Application US/10374979
; Publication No. US20030219793A1
; GENERAL INFORMATION:
; APPLICANT: John P. Carulli et al.
; TITLE OF INVENTION: THE HIGH BONE MASS GENE OF 11q13.3
; FILE REFERENCE: 032796-021
; CURRENT APPLICATION NUMBER: US/10/374,979
; PRIOR FILING DATE: 2003-03-04
; PRIOR APPLICATION NUMBER: US 09/544,398
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/543,771
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 109
; SEQ ID NO 10
; LENGTH: 8705
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-374-979-10

Query Match      1.6%; Score 50; DB 6; Length 8705;
Best Local Similarity 100.0%; Pred. No. 5.3e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122
DB      6492 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 6541

RESULT 77
US-10-182-936A-10
; Sequence 10, Application US/10182936A
; Publication No. US20040038660A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Kristina M.
; APPLICANT: Anisowicz, Anthony
; APPLICANT: Bhat, Bhenn
; APPLICANT: Damagnez, Veronique
; APPLICANT: Robinson, John
; APPLICANT: Yaworsky, Paul
; TITLE OF INVENTION: Reagents and Method for Modulating DKK-Mediated Interactions
; FILE REFERENCE: 032796-143
; CURRENT APPLICATION NUMBER: US/10/182,936A
; PRIOR FILING DATE: 2002-08-02
; PRIOR APPLICATION NUMBER: PCT/US02/15982
; PRIOR FILING DATE: 2002-05-17
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
```

```
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 216
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 8705
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-182-936A-10

Query Match      1.6%; Score 50; DB 7; Length 8705;
Best Local Similarity 100.0%; Pred. No. 5.3e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122
DB      6492 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 6541

RESULT 78
US-10-731-739-10
; Sequence 10, Application US/10731739
; Publication No. US20040176582A1
; GENERAL INFORMATION:
; APPLICANT: Carulli, John P.
; APPLICANT: Little, Randall D.
; APPLICANT: Becker, Robert R.
; APPLICANT: Johnson, Mark L.
; TITLE OF INVENTION: High bone mass gene of 11q13.3
; FILE REFERENCE: 032796-013
; CURRENT APPLICATION NUMBER: US/10/731,739
; PRIOR FILING DATE: 2003-12-10
; PRIOR APPLICATION NUMBER: US/09/544,398B
; PRIOR FILING DATE: 2002-06-10
; PRIOR APPLICATION NUMBER: US 09/229,319
; PRIOR FILING DATE: 1999-01-13
; PRIOR APPLICATION NUMBER: US 60/071,449
; PRIOR FILING DATE: 1998-01-13
; PRIOR APPLICATION NUMBER: US 60/105,511
; PRIOR FILING DATE: 1998-10-23
; NUMBER OF SEQ ID NOS: 641
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 8705
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-731-739-10

Query Match      1.6%; Score 50; DB 8; Length 8705;
Best Local Similarity 100.0%; Pred. No. 5.3e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122
DB      6492 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 6541

RESULT 79
US-10-477-238A-10
; Sequence 10, Application US/10477238A
; Publication No. US2004022136A1
; GENERAL INFORMATION:
; APPLICANT: Babi, Philip
; APPLICANT: Yaworsky, Paul
; APPLICANT: Bex, Frederick J. III
; APPLICANT: Bodine, Peter Van Ness
; TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
; FILE REFERENCE: 032796-212
; CURRENT APPLICATION NUMBER: US/10/477,238A
; PRIOR FILING DATE: 2003-11-10
; PRIOR APPLICATION NUMBER: US 60/290,071
```

```

? PRIOR FILING DATE: 2001-05-11
? PRIOR APPLICATION NUMBER: US 60/291,311
? PRIOR FILING DATE: 2001-05-17
? PRIOR APPLICATION NUMBER: US 60/353,058
? PRIOR FILING DATE: 2002-02-01
? PRIOR APPLICATION NUMBER: US 60/361,293
? PRIOR FILING DATE: 2002-03-04
? NUMBER OF SEQ ID NOS: 812
? SOFTWARE: FastSeq for Windows Version 4.0
? SEQ ID NO 10
? LENGTH: 8705
? TYPE: DNA
? ORGANISM: Homo sapiens
US-10-477-238A-10

```

Query Match	1.6%	Score 50;	DB 8;	Length 8705;
Best Local Similarity	100.0%	Pred. No. 5.3e-15;		
Matches 50;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

```

Oy      3073 AGATTGTGCCACTGCACCTCCAGCCTGGGGCAACAGAGCAACTCTGTCTC 31222
          |||||
Db      6492 AGATTGTGCCACTGCACCTCCAGCCTGGGGCAACAGAGCAACTCTGTCTC 65411

```

```

RESULT 80
US-10-680-287A-10
; Sequence 10, Application US/10680287A
; Publication No. US20040244069A1
; GENERAL INFORMATION:
; APPLICANT: Bablitz, Philip
; APPLICANT: Yaworsky, Paul
; APPLICANT: Bex, Frederick J. III
; APPLICANT: Bodine, Peter Van Nest
; TITLE OF INVENTION: Transgenic Animal Model of Bone Mass Modulation
; FILE REFERENCE: 032796-179
; CURRENT APPLICATION NUMBER: US/10/680,287A
; PRIOR FILING DATE: 2003-10-08
; PRIOR APPLICATION NUMBER: PCT/US02/14876
; PRIOR FILING DATE: 2002-05-13
; PRIOR APPLICATION NUMBER: US 60/290,071
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US 60/291,311
; PRIOR FILING DATE: 2001-05-17
; PRIOR APPLICATION NUMBER: US 60/353,058
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: US 60/361,293
; PRIOR FILING DATE: 2002-03-04
; NUMBER OF SEQ ID NOS: 812
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 8705
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-680-287A-10

```

Query Match	1.6%;	Score 50;	DB 8;	Length 8705;
Best Local Similarity	100.0%;	Pred. No. 5.3e-15;		
Matches	50;	Conservative	0;	Mismatches 0;
				Indels 0;
				Gaps 0

Qy	3073	AGATTGTCAC	TGCATCTC	CAGCCCTGG	CAACAGAC	CAAGACTCTG	CTC	3122
Db	6492	AGATTGTCAC <td>TGCATCTC <td>CAGCCCTGG <td>CAACAGAC <td>CAAGACTCTG <td>CTC</td> <td>6541</td> </td></td></td></td>	TGCATCTC <td>CAGCCCTGG <td>CAACAGAC <td>CAAGACTCTG <td>CTC</td> <td>6541</td> </td></td></td>	CAGCCCTGG <td>CAACAGAC <td>CAAGACTCTG <td>CTC</td> <td>6541</td> </td></td>	CAACAGAC <td>CAAGACTCTG <td>CTC</td> <td>6541</td> </td>	CAAGACTCTG <td>CTC</td> <td>6541</td>	CTC	6541

RESULT 81  
US-10-477-173-10  
; Sequence 10, Application US/10477173  
; Publication No. US20050070659A1  
; GENERAL INFORMATION:  
; APPLICANT: Genome Therapeutics Corporation and  
; APPLICANT: Allen, Kristina M.  
; APPLICANT: Yaworsky, Paul  
; APPLICANT: Morales, Arturo J.

```

1  APPLICANT: Graham, James R.
2  APPLICANT: Antosiewicz, Anthony
3  APPLICANT: Liu, Wei
4  TITLE OF INVENTION: HBM Variants that Modulate Bone Mass and Lipid Levels
5  FILE REFERENCE: 032796-135
6  CURRENT APPLICATION NUMBER: US/10/477,173
7  CURRENT FILING DATE: 2003-11-10
8  PRIOR APPLICATION NUMBER: US 60/290,071
9  PRIOR FILING DATE: 2001-05-11
10 PRIOR APPLICATION NUMBER: US 60/291,311
11 PRIOR FILING DATE: 2001-05-17
12 PRIOR APPLICATION NUMBER: US 60/353,058
13 PRIOR FILING DATE: 2002-02-01
14 PRIOR APPLICATION NUMBER: US 60/361,293
15 PRIOR FILING DATE: 2002-03-04
16 NUMBER OF SEQ ID NOS: 1086
17 SOFTWARE: FastSeq for Windows Version 4.0
18 SEQ ID NO 10
19 LENGTH: 8705
20 TYPE: DNA
21 ORGANISM: Homo sapiens
22 US-10-477-173-10

```

Query Match	1.6%	Score 50;	DB 9;	Length 8705;
Best Local Similarity	100.0%	Pred. No. 5.3e-15;		
Matches 50; Conservative	0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy 3073 AGATTGTGCACTGCACCTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122

Db 6492 AGATTGTGCACTGCACCTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 6541

```

RESULT 82
US-10-834-377-10
# Sequence 10, Application US/10834377
# Publication No. US20050142617A1
# GENERAL INFORMATION:
# APPLICANT: Carrull, John P.
# APPLICANT: Little, Randall D.
# APPLICANT: Recker, Robert R.
# APPLICANT: Johnson, Mark L.
# TITLE OF INVENTION: High bone mass gene of 11q13.3
# FILE REFERENCE: 032796-04
# CURRENT APPLICATION NUMBER: US/10/834,377
# CURRENT FILING DATE: 2004-04-29
# PRIOR APPLICATION NUMBER: US/09/543,771B
# PRIOR FILING DATE: 2000-04-05
# PRIOR APPLICATION NUMBER: US 09/229,319
# PRIOR FILING DATE: 1999-01-13
# PRIOR APPLICATION NUMBER: US 60/071,449
# PRIOR FILING DATE: 1998-01-13
# PRIOR APPLICATION NUMBER: US 60/105,511
# PRIOR FILING DATE: 1998-10-23
# NUMBER OF SEQ ID NOS: 641
# SOFTWARE: FastSeq for Windows Version 4.0
# SEQ ID NO 10
# LENGTH: 8705
# TYPE: DNA
# ORGANISM: Homo sapiens
# US-10-834-377-10

```

Query Match	1.6%	Score 50;	DB 9;	Length 8705;
Best Local Similarity	100.0%	Pred. No. 5.3e-15;		
Matches 50;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	3073	AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC	3122
Db	6492	AGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC	6541

RESULT 83  
US-09-764-891-8149/c  
; Sequence 8149, Application US/09764891

```
Publication No. US20030077808A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC006
CURRENT APPLICATION NUMBER: US/09/764,891
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 10231
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 8149
LENGTH: 31474
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-891-8149

Query Match      1.6%; Score 50; DB 3; Length 31474;
Best Local Similarity 100.0%; Pred. No. 5.1e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122
DB      72 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 23

RESULT 84
US-09-764-878-379/c
Sequence 379, Application US/09764878
Patent No. US2002009015A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PAL21
CURRENT APPLICATION NUMBER: US/09/764,878
CURRENT FILING DATE: 2001-01-17
Prior application data removed - consult PALM or file wrapper
NUMBER OF SEQ ID NOS: 428
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 379
LENGTH: 32189
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-878-379

Query Match      1.6%; Score 50; DB 3; Length 32189;
Best Local Similarity 100.0%; Pred. No. 5.1e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122
DB      12517 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 12468

RESULT 85
US-10-079-854-379/c
Sequence 379, Application US/10079854
Publication No. US20030054368A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PAL21C1
CURRENT APPLICATION NUMBER: US/10/079,854
CURRENT FILING DATE: 2002-02-22
Prior application removed - See file wrapper or Palm
NUMBER OF SEQ ID NOS: 428
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 379
LENGTH: 32189
TYPE: DNA
ORGANISM: Homo sapiens
US-10-079-854-379

Query Match      1.6%; Score 50; DB 5; Length 32189;
```

```
Best Local Similarity 100.0%; Pred. No. 5.1e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122
DB      12517 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 12468

RESULT 86
US-09-764-877-2623/c
Sequence 2623, Application US/09764877
Patent No. US20020147140A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PC005
CURRENT APPLICATION NUMBER: US/09/764,877
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 4031
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 2623
LENGTH: 32193
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-877-2623

Query Match      1.6%; Score 50; DB 3; Length 32193;
Best Local Similarity 100.0%; Pred. No. 5.1e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122
DB      28855 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTCTC 28806

RESULT 87
US-10-216-464-37/c
Sequence 37, Application US/10216464
Publication No. US20030207285A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT219C1
CURRENT APPLICATION NUMBER: US/10/216,464
CURRENT FILING DATE: 2002-08-12
PRIOR APPLICATION NUMBER: 09/764,883
PRIOR FILING DATE: 2001-01-17
PRIOR APPLICATION NUMBER: 60/179,065
PRIOR FILING DATE: 2000-01-31
PRIOR APPLICATION NUMBER: 60/180,628
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/214,886
PRIOR FILING DATE: 2000-06-28
PRIOR APPLICATION NUMBER: 60/217,487
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,758
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/220,963
PRIOR FILING DATE: 2000-07-26
PRIOR APPLICATION NUMBER: 60/217,496
PRIOR FILING DATE: 2000-07-11
PRIOR APPLICATION NUMBER: 60/225,447
PRIOR FILING DATE: 2000-08-14
PRIOR APPLICATION NUMBER: 60/218,290
PRIOR FILING DATE: 2000-07-14
Remaining Prior Application data removed - See file wrapper or PALM.
NUMBER OF SEQ ID NOS: 38
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 37
LENGTH: 32193
TYPE: DNA
ORGANISM: Homo sapiens
```

US-10-216-464-37

Query Match 1.6%; Score 50; DB 6; Length 32193;  
Best Local Similarity 100.0%; Pred. No. 5.1e-15;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACGAACTGTCTC 3122  
DB 28855 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACGAACTGTCTC 28806

RESULT 88  
US-10-242-515-2623/c  
; Sequence 2623, Application US/10242515  
; Publication No. US20040009488A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC005C1  
; CURRENT APPLICATION NUMBER: US/10/242,515  
; CURRENT FILING DATE: 2002-09-13  
; PRIOR APPLICATION NUMBER: 09/764,877  
; PRIOR FILING DATE: 2001-01-17  
; PRIOR APPLICATION NUMBER: 60/179,065  
; PRIOR FILING DATE: 2000-01-31  
; PRIOR APPLICATION NUMBER: 60/180,628  
; PRIOR FILING DATE: 2000-02-04  
; PRIOR APPLICATION NUMBER: 60/214,886  
; PRIOR FILING DATE: 2000-06-28  
; PRIOR APPLICATION NUMBER: 60/217,487  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,758  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/220,963  
; PRIOR FILING DATE: 2000-07-26  
; PRIOR APPLICATION NUMBER: 60/217,496  
; PRIOR FILING DATE: 2000-07-11  
; PRIOR APPLICATION NUMBER: 60/225,447  
; PRIOR FILING DATE: 2000-08-14  
; PRIOR APPLICATION NUMBER: 60/218,290  
; PRIOR FILING DATE: 2000-07-14  
; Remaining Prior Application data removed - See File Wrapper or PALM.  
; NUMBER OF SEQ ID NOS: 4031  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 2623  
; LENGTH: 32193  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-242-515-2623

Query Match 1.6%; Score 50; DB 6; Length 32193;  
Best Local Similarity 100.0%; Pred. No. 5.1e-15;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACGAACTGTCTC 3122  
DB 28855 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACGAACTGTCTC 28806

RESULT 89  
US-09-764-878-377/c  
; Sequence 377, Application US/09764878  
; Patent No. US20020090615A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA121  
; CURRENT APPLICATION NUMBER: US/09/764,878  
; CURRENT FILING DATE: 2001-01-17  
; Prior Application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 428  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 377

; LENGTH: 32221  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (7464)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-09-764-878-377

Query Match 1.6%; Score 50; DB 3; Length 32221;  
Best Local Similarity 100.0%; Pred. No. 5.1e-15;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACGAACTGTCTC 3122  
DB 12548 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACGAACTGTCTC 12499

RESULT 90  
US-10-079-854-377/c  
; Sequence 377, Application US/10079854  
; Publication No. US20030054368A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PA121C1  
; CURRENT APPLICATION NUMBER: US/10/079,854  
; CURRENT FILING DATE: 2002-02-22  
; Prior Application removed - See File Wrapper or Palm  
; NUMBER OF SEQ ID NOS: 428  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 377  
; LENGTH: 32221  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (7464)  
; OTHER INFORMATION: n equals a,t,g, or c  
US-10-079-854-377

Query Match 1.6%; Score 50; DB 5; Length 32221;  
Best Local Similarity 100.0%; Pred. No. 5.1e-15;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACGAACTGTCTC 3122  
DB 12548 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACGAACTGTCTC 12499

RESULT 91  
US-10-085-117-190  
; Sequence 190, Application US/10085117  
; Publication No. US2003023234A1  
; GENERAL INFORMATION:  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER  
; FILE REFERENCE: 529452000121  
; CURRENT APPLICATION NUMBER: US/10/085,117  
; CURRENT FILING DATE: 2002-02-27  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 361  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 190  
; LENGTH: 33126  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-085-117-190

Query Match 1.6%; Score 50; DB 6; Length 33126;  
Best Local Similarity 100.0%; Pred. No. 5.1e-15;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGCAAGACTGTGTCTC 3122

DB 30352 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGCAAGACTGTGTCTC 30401

## RESULT 92

US-10-741-600-17637  
; Sequence 17637, Application US/10741600  
; Publication No. US20050026169A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CL001499  
; CURRENT APPLICATION NUMBER: US/10/741,600  
; PRIOR FILING DATE: 2003-12-22  
; NUMBER OF SEQ ID NOS: 73997  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 17637  
; LENGTH: 56510  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-741-600-17637

Query Match 1.6%; Score 50; DB 8; Length 56510;  
Best Local Similarity 100.0%; Pred. No. 5e-15;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCGAGTGTGATCCTGAGGCGAGGAGTTGAGAGCCCTGGCCCAAC 2938

DB 4790 GAGGCGAGTGTGATCCTGAGGCGAGGAGTTGAGAGCCCTGGCCCAAC 4839

## RESULT 93

US-10-087-192-574/C  
; Sequence 574, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 574  
; LENGTH: 66973  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(66973)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-087-192-574

Query Match 1.6%; Score 50; DB 5; Length 66973;  
Best Local Similarity 100.0%; Pred. No. 5e-15;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGCAAGACTGTGTCTC 3122

DB 29006 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGCAAGACTGTGTCTC 28957

## RESULT 94

US-10-087-192-1408

; Sequence 1408, Application US/10087192  
; Publication No. US20020182586A1  
; GENERAL INFORMATION:  
; APPLICANT: Morris, David W.  
; APPLICANT: Engelhard, Eric K.  
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR  
; FILE REFERENCE: 529452000122  
; CURRENT APPLICATION NUMBER: US/10/087,192  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 09/747,377  
; PRIOR FILING DATE: 2000-12-22  
; PRIOR APPLICATION NUMBER: US 09/798,586  
; PRIOR FILING DATE: 2001-03-02  
; NUMBER OF SEQ ID NOS: 2059  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 1408  
; LENGTH: 156843  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc\_feature  
; LOCATION: (1)...(156843)  
; OTHER INFORMATION: n = A,T,C or G  
US-10-087-192-1408

Query Match 1.6%; Score 50; DB 5; Length 156843;  
Best Local Similarity 100.0%; Pred. No. 4.9e-15;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGCAAGACTGTGTCTC 3122

DB 148232 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGCAAGACTGTGTCTC 148281

## RESULT 95

US-10-719-993-6812  
; Sequence 6812, Application US/10719993  
; Publication No. US20040265849A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CL001496  
; CURRENT APPLICATION NUMBER: US/10/719,993  
; PRIOR FILING DATE: 2003-11-24  
; NUMBER OF SEQ ID NOS: 55342  
; SOFTWARE: FASTSEQ for Windows Version 4.0  
; SEQ ID NO 6812  
; LENGTH: 202814  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-719-993-6812

Query Match 1.6%; Score 50; DB 8; Length 202814;  
Best Local Similarity 100.0%; Pred. No. 4.9e-15;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGCAAGACTGTGTCTC 3122

DB 145090 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGCAAGACTGTGTCTC 145139

## RESULT 96

US-10-271-416-9  
; Sequence 9, Application US/10271416  
; Publication No. US20040043021A1  
; GENERAL INFORMATION:  
; APPLICANT: Keith, Tim  
; APPLICANT: Little, Randall D.  
; APPLICANT: Van Berdewegh, Paul  
; APPLICANT: Dupuis, Josee

```
/ APPLICANT: Del Maestro, Richard G.
/ APPLICANT: Allen, Kristina
/ TITLE OF INVENTION: NUCLEOTIDE AND AMINO ACID SEQUENCES
/ TITLE OF INVENTION: RELATING TO RESPIRATORY DISEASES AND OBESITY
/ FILE REFERENCE: 2976-4045
/ CURRENT APPLICATION NUMBER: US/10/271,416
/ PRIOR FILING DATE: 2002-10-11
/ PRIOR APPLICATION NUMBER: 60/328,424
/ PRIOR FILING DATE: 2001-10-11
/ NUMBER OF SEQ ID NOS: 9
/ SOFTWARE: FaastSeq for Windows Version 4.0
/ SEQ ID NO 9
/ LENGTH: 276820
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: conflict
/ LOCATION: (167043) ... (167043)
/ OTHER INFORMATION: Y=C or T
US-10-271-416-9
```

```
Query Match 1.6%; Score 50; DB 7; Length 276820;
Best Local Similarity 100.0%; Pred. No. 4,8e-15;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3073 AGATTGCGCACTGCACTCCAGCCTGGGCAAGAGCAAGACTCTGTCTC 3122
DB 194263 AGATTGCGCACTGCACTCCAGCCTGGGCAAGAGCAAGACTCTGTCTC 194312
```

```
RESULT 97
US-09-925-065A-140670/c
/ Sequence 140670, Application US/09925065A
/ Publication No. US20050228172A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-05-09
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FaastSeq for Windows Version 4.0
/ SEQ ID NO 140670
/ LENGTH: 420
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-140670
```

```
Query Match 1.6%; Score 49; DB 4; Length 420;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3074 GATTGCGCACTGCACTCCAGCCTGGGCAAGAGCAAGACTCTGTCTC 3122
DB 273 GATTGCGCACTGCACTCCAGCCTGGGCAAGAGCAAGACTCTGTCTC 225
```

```
RESULT 98
US-09-925-065A-140671/c
/ Sequence 140671, Application US/09925065A
/ Publication No. US20050228172A9
/ GENERAL INFORMATION:
```

```
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FaastSeq for Windows Version 4.0
/ SEQ ID NO 140671
/ LENGTH: 532
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-140671
```

```
Query Match 1.6%; Score 49; DB 4; Length 532;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3074 GATTGCGCACTGCACTCCAGCCTGGGCAAGAGCAAGACTCTGTCTC 3122
DB 232 GATTGCGCACTGCACTCCAGCCTGGGCAAGAGCAAGACTCTGTCTC 184
```

```
RESULT 99
US-09-925-065A-18609/c
/ Sequence 18609, Application US/09925065A
/ Publication No. US20050228172A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-05-09
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FaastSeq for Windows Version 4.0
/ SEQ ID NO 18609
/ LENGTH: 552
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-18609
```

```
Query Match 1.6%; Score 49; DB 4; Length 552;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3073 AGATTGCGCACTGCACTCCAGCCTGGGCAAGAGCAAGACTCTGTCTC 3121
DB 363 AGATTGCGCACTGCACTCCAGCCTGGGCAAGAGCAAGACTCTGTCTC 315
```

```
RESULT 100
US-09-925-065A-18610/c
```

```
/ Sequence 18610, Application US/09925065A
/ Publication No. US20050228172A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 18610
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-18610
```

```
Query Match 1.6%; Score 49; DB 4; Length 552;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3073 AGATTGTGCACCTGCCTCCAGCTGGGCAACAGACGACTCTGTCT 3121
DB 363 AGATTGTGCACCTGCCTCCAGCTGGGCAACAGACGACTCTGTCT 315
```

```
RESULT 101
US-09-925-065A-18611/C
/ Sequence 18611, Application US/09925065A
/ Publication No. US20050228172A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 18611
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-18611
```

```
Query Match 1.6%; Score 49; DB 4; Length 552;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3073 AGATTGTGCACCTGCCTCCAGCTGGGCAACAGACGACTCTGTCT 3121
DB 363 AGATTGTGCACCTGCCTCCAGCTGGGCAACAGACGACTCTGTCT 315
```

```
RESULT 102
US-10-027-632-80386/C
/ Sequence 80386, Application US/10027632
/ Publication No. US20020198371A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 80386
/ LENGTH: 580
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(580)
/ OTHER INFORMATION: n = A,T,C or G
US-10-027-632-80386
```

```
Query Match 1.6%; Score 49; DB 5; Length 580;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3074 GATTGTGCACCTGCCTCCAGCTGGGCAACAGACGACTCTGTCT 3122
DB 281 GATTGTGCACCTGCCTCCAGCTGGGCAACAGACGACTCTGTCT 233
```

```
RESULT 103
US-10-027-632-80387/C
/ Sequence 80387, Application US/10027632
/ Publication No. US20020198371A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
```



```
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 80387
/ LENGTH: 580
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(580)
/ OTHER INFORMATION: n = A,T,C or G
US-10-027-632-80387

Query Match      1.6%; Score 49; DB 5; Length 580;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3074 GATTGTGCCACTGCCTCCAGCTGGGCAACAGACCAAGACTCTGTCTC 3122
Db      281 GATTGTGCCACTGCCTCCAGCTGGGCAACAGACCAAGACTCTGTCTC 233

RESULT 104
US-10-027-632-80386/c
/ Sequence 80386, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ POLYMORPHISMS IN THE HUMAN GENOME
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 80386
/ LENGTH: 580
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(580)
/ OTHER INFORMATION: n = A,T,C or G
US-10-027-632-80386

Query Match      1.6%; Score 49; DB 6; Length 580;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3074 GATTGTGCCACTGCCTCCAGCTGGGCAACAGACCAAGACTCTGTCTC 3122
Db      281 GATTGTGCCACTGCCTCCAGCTGGGCAACAGACCAAGACTCTGTCTC 233

RESULT 105
US-10-027-632-80387/c
/ Sequence 80387, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
```

```
/ TITLE OF INVENTION: Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 80387
/ LENGTH: 580
/ TYPE: DNA
/ ORGANISM: Human
/ FEATURE:
/ NAME/KEY: misc_feature
/ LOCATION: (1)...(580)
/ OTHER INFORMATION: n = A,T,C or G
US-10-027-632-80387

Query Match      1.6%; Score 49; DB 6; Length 580;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3074 GATTGTGCCACTGCCTCCAGCTGGGCAACAGACCAAGACTCTGTCTC 3122
Db      281 GATTGTGCCACTGCCTCCAGCTGGGCAACAGACCAAGACTCTGTCTC 233

RESULT 106
US-09-925-065A-708700/c
/ Sequence 708700, Application US/09925065A
/ Publication No. US20050228172A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ NUCLEOTIDE POLYMORPHISMS IN THE HUMAN GENOME
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 708700
/ LENGTH: 695
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ US-09-925-065A-708700

Query Match      1.6%; Score 49; DB 4; Length 695;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2892 GCAGGTGATCCTGAGGCCAGAGTTCGAGACCAAGCCTGGCCAACAT 2940
```

Db 487 GAGGTGGATCACTGAGGCCAGAGATTGAGACCAAGCTGGCCAAAT 439

## RESULT 107

US-09-925-065A-932035/C  
Sequence 932035, Application US/09925065A  
Publication No. US20050228172A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single  
FILE REFERENCE: 108827.135  
CURRENT APPLICATION NUMBER: US/09/925,065A  
CURRENT FILING DATE: 2001-08-08  
PRIOR APPLICATION NUMBER: US 60/243,096  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: US 60/252,147  
PRIOR FILING DATE: 2000-11-20  
PRIOR APPLICATION NUMBER: US 60/250,092  
PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: US 60/261,766  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/289,846  
PRIOR FILING DATE: 2001-05-09  
NUMBER OF SEQ ID NOS: 957086  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 932035  
LENGTH: 750  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-925-065A-932035

Query Match 1.6%; Score 49; DB 4; Length 750;  
Best Local Similarity 100.0%; Pred. No. 1.9e-14;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3074 GATTGGCCACTGCAGCTCCAGCTGGGCAACAGAGAACTGTCTC 3122  
DB 612 GATTGGCCACTGCAGCTCCAGCTGGGCAACAGAGAACTGTCTC 564

## RESULT 108

US-09-925-065A-932036/C  
Sequence 932036, Application US/09925065A  
Publication No. US20050228172A9  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single  
FILE REFERENCE: 108827.135  
CURRENT APPLICATION NUMBER: US/09/925,065A  
CURRENT FILING DATE: 2001-08-08  
PRIOR APPLICATION NUMBER: US 60/243,096  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: US 60/252,147  
PRIOR FILING DATE: 2000-11-20  
PRIOR APPLICATION NUMBER: US 60/250,092  
PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: US 60/261,766  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/289,846  
PRIOR FILING DATE: 2001-05-09  
NUMBER OF SEQ ID NOS: 957086  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 932036  
LENGTH: 750  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-925-065A-932036

Query Match 1.6%; Score 49; DB 4; Length 750;  
Best Local Similarity 100.0%; Pred. No. 1.9e-14;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3074 GATTGGCCACTGCAGCTCCAGCTGGGCAACAGAGAACTGTCTC 3122  
DB 612 GATTGGCCACTGCAGCTCCAGCTGGGCAACAGAGAACTGTCTC 564

## RESULT 109

US-10-027-632-113573  
Sequence 113573, Application US/10027632  
Publication No. US20020198371A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002  
PRIOR FILING DATE: 1999-08-09  
NUMBER OF SEQ ID NOS: 325720  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 113573  
LENGTH: 3170  
TYPE: DNA  
ORGANISM: Human  
US-10-027-632-113573

Query Match 1.6%; Score 49; DB 5; Length 3170;  
Best Local Similarity 100.0%; Pred. No. 1.8e-14;  
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGGCCACTGCAGCTCCAGCTGGGCAACAGAGAACTGTCTC 3121  
DB 1985 AGATTGGCCACTGCAGCTCCAGCTGGGCAACAGAGAACTGTCTC 2033

## RESULT 110

US-10-027-632-113574  
Sequence 113574, Application US/10027632  
Publication No. US20020198371A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
FILE REFERENCE: 108827.129  
CURRENT APPLICATION NUMBER: US/10/027,632  
CURRENT FILING DATE: 2002-04-30  
PRIOR APPLICATION NUMBER: US 60/218,006  
PRIOR FILING DATE: 2000-07-12  
PRIOR APPLICATION NUMBER: US 60/198,676  
PRIOR FILING DATE: 2000-04-20  
PRIOR APPLICATION NUMBER: US 60/193,483  
PRIOR FILING DATE: 2000-03-29  
PRIOR APPLICATION NUMBER: US 60/185,218  
PRIOR FILING DATE: 2000-02-24  
PRIOR APPLICATION NUMBER: US 60/167,363  
PRIOR FILING DATE: 1999-11-23  
PRIOR APPLICATION NUMBER: US 60/156,358  
PRIOR FILING DATE: 1999-09-28  
PRIOR APPLICATION NUMBER: US 60/146,002

```
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 113574
/ LENGTH: 3170
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-113574

Query Match
Best Local Similarity 1.6%; Score 49; DB 5; Length 3170;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAAGAGACGACTGTCTCT 3121
Db 1985 AGATTGTGCACCTGCACCTCCAGCTGGGCAAGAGACGACTGTCTCT 2033

RESULT 111
US-10-027-632-113573
/ Sequence 113573, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 113573
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-113573

Query Match
Best Local Similarity 1.6%; Score 49; DB 6; Length 3170;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAAGAGACGACTGTCTCT 3121
Db 1985 AGATTGTGCACCTGCACCTCCAGCTGGGCAAGAGACGACTGTCTCT 2033

RESULT 112
US-10-027-632-113574
/ Sequence 113574, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
```

```
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 113574
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-113574

Query Match
Best Local Similarity 1.6%; Score 49; DB 6; Length 3170;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAAGAGACGACTGTCTCT 3121
Db 1985 AGATTGTGCACCTGCACCTCCAGCTGGGCAAGAGACGACTGTCTCT 2033

RESULT 113
US-10-027-632-182264
/ Sequence 182264, Application US/10027632
/ Publication No. US20020198371A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ PRIOR FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 182264
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-182264

Query Match
Best Local Similarity 1.5%; Score 48; DB 5; Length 460;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3075 ATTGTGCACCTGCACCTCCAGCTGGGCAAGAGACGACTGTCTCTC 3122
Db 160 ATTGTGCACCTGCACCTCCAGCTGGGCAAGAGACGACTGTCTCTC 207

RESULT 114
US-10-027-632-319423
```

```
; Sequence 319423, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 319423
; LENGTH: 460
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-319423
```

```
Query Match      1.5%; Score 48; DB 5; Length 460;
Best Local Similarity 100.0%; Pred. No. 6,4e-14;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3075 ATTGTCCACTGCATCTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
DB      160 ATTGTCCACTGCATCTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 207
```

```
RESULT 115
US-10-027-632-182264
; Sequence 182264, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 182264
; LENGTH: 460
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-182264
```

```
Query Match      1.5%; Score 48; DB 6; Length 460;
Best Local Similarity 100.0%; Pred. No. 6,4e-14;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3075 ATTGTCCACTGCATCTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
DB      160 ATTGTCCACTGCATCTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 207
```

```
RESULT 116
US-10-027-632-319423
; Sequence 319423, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 319423
; LENGTH: 460
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-319423
```

```
Query Match      1.5%; Score 48; DB 6; Length 460;
Best Local Similarity 100.0%; Pred. No. 6,4e-14;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3075 ATTGTCCACTGCATCTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
DB      160 ATTGTCCACTGCATCTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 207
```

```
RESULT 117
US-10-027-632-132626
; Sequence 132626, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
```

```

; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 132626
; LENGTH: 716
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-132626
```

```

Query Match          1.5%; Score 48; DB 5; Length 716;
Best Local Similarity 100.0%; Pred. No. 6.4e-14;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3075 ATTGTGCCACTGCATCCGACCTGGGCAACAGAGCAAGACTGTCTC 3122
Db      358 ATTGTGCCACTGCATCCGACCTGGGCAACAGAGCAAGACTGTCTC 405
```

RESULT 118

```

US-10-027-632-132626
; Sequence 132626, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
```

APPLICANT: Mang, David G.

TITLE OF INVENTION: Identification and Mapping of Single Nucleotide

FILE REFERENCE: 108827.129

CURRENT APPLICATION NUMBER: US/10/027,632

CURRENT FILING DATE: 2002-04-30

PRIOR APPLICATION NUMBER: US 60/218,006

PRIOR FILING DATE: 2000-07-12

PRIOR APPLICATION NUMBER: US 60/198,676

PRIOR FILING DATE: 2000-04-20

PRIOR APPLICATION NUMBER: US 60/193,483

PRIOR FILING DATE: 2000-03-29

PRIOR APPLICATION NUMBER: US 60/185,218

PRIOR FILING DATE: 2000-02-24

PRIOR APPLICATION NUMBER: US 60/167,363

PRIOR FILING DATE: 1999-11-23

PRIOR APPLICATION NUMBER: US 60/156,358

PRIOR FILING DATE: 1999-09-28

PRIOR APPLICATION NUMBER: US 60/146,002

PRIOR FILING DATE: 1999-08-09

NUMBER OF SEQ ID NOS: 325720

SOFTWARE: FastSeq for Windows Version 4.0

SEQ ID NO 132626

LENGTH: 716

TYPE: DNA

ORGANISM: Human

US-10-027-632-132626

Query Match 1.5%; Score 48; DB 6; Length 716;

Best Local Similarity 100.0%; Pred. No. 6.4e-14;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      3075 ATTGTGCCACTGCATCCGACCTGGGCAACAGAGCAAGACTGTCTC 3122
Db      358 ATTGTGCCACTGCATCCGACCTGGGCAACAGAGCAAGACTGTCTC 405
```

RESULT 119

```

US-10-052-482-172
; Sequence 172, Application US/10052482
; Publication No. US20040072264A1
; GENERAL INFORMATION:
```

APPLICANT: Engelhard, Eric

TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER

FILE REFERENCE: A-71087/RMS/DCP

CURRENT APPLICATION NUMBER: US/10/052,482

CURRENT FILING DATE: 2002-08-15

```

; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 172
; LENGTH: 52242
; TYPE: DNA
; ORGANISM: Homo sapiens
```

FEATURE: NAME/KEY: misc.feature

LOCATION: (51242)..(51261)

OTHER INFORMATION: "n" at positions 51242 to 51261 can be any base

US-10-052-482-172

```

Query Match          1.5%; Score 48; DB 7; Length 52242;
Best Local Similarity 100.0%; Pred. No. 5.7e-14;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2895 GGTGATCACCCTGAGGCCAGAGTTGAGACCGCTGGCCAACTAG 2942
Db      5520 GGTGATCACCCTGAGGCCAGAGTTGAGACCGCTGGCCAACTAG 5567
```

RESULT 120

US-10-737-082-85/C

Sequence 85, Application US/10737082

Publication No. US20050130170A1

GENERAL INFORMATION:

APPLICANT: Bayer Healthcare LLC

APPLICANT: Beard, Chris

APPLICANT: Burgess, Chris

APPLICANT: Gannon, Allison

APPLICANT: Harvey, Jeanne

APPLICANT: Lechner, John F.

APPLICANT: Li, Zheng

TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences

FILE REFERENCE: 1657/2032

CURRENT APPLICATION NUMBER: US/10/737,082

CURRENT FILING DATE: 2003-12-16

PRIOR APPLICATION NUMBER: US 10/737,082

PRIOR FILING DATE: 2003-12-16

NUMBER OF SEQ ID NOS: 300

SOFTWARE: PatentIn version 3.2

SEQ ID NO 85

LENGTH: 53779

TYPE: DNA

ORGANISM: Homo sapiens

US-10-737-082-85

Query Match 1.5%; Score 48; DB 9; Length 53779;

Best Local Similarity 100.0%; Pred. No. 5.7e-14;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY      2889 GAGGAGGTGATCCTGAGGCCAGAGTTGAGACCGCTGGCCA 2936
Db      2311 GAGGAGGTGATCCTGAGGCCAGAGTTGAGACCGCTGGCCA 2264
```

RESULT 121

```

US-10-765-790-85/C
; Sequence 85, Application US/10765790
; Publication No. US20050130172A1
; GENERAL INFORMATION:
```

APPLICANT: Bayer Healthcare LLC

APPLICANT: Beard, Chris

APPLICANT: Burgess, Chris

APPLICANT: Gannon, Allison

APPLICANT: Harvey, Jeanne

APPLICANT: Lechner, John F.

APPLICANT: Li, Zheng

TITLE OF INVENTION: Identification and Verification of Methylation Marker Sequences

```
FILE REFERENCE: 1657/2035
CURRENT APPLICATION NUMBER: US/10/765,790
CURRENT FILING DATE: 2004-01-27
PRIOR APPLICATION NUMBER: US 10/737,082
PRIOR FILING DATE: 2003-12-16
NUMBER OF SEQ ID NOS: 300
SOFTWARE: Patentin version 3.2
SEQ ID NO 85
LENGTH: 53779
TYPE: DNA
ORGANISM: Homo sapiens
US-10-765-790-85

Query Match
Best Local Similarity 100.0%; Score 48; DB 9; Length 53779;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2889 GAGGCAAGTGTGATCACTGAGGCGAGAGTTCGAGACGAGCTGGCCA 2936
Db 2311 GAGGCAAGTGTGATCACTGAGGCGAGAGTTCGAGACGAGCTGGCCA 2264

RESULT 122
US-10-741-600-17651
Sequence 17651, Application US/10741600
Publication No. US20050026169A1
GENERAL INFORMATION:
APPLICANT: CARGILL, Michele et al.
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
FILE REFERENCES: CL001439
CURRENT APPLICATION NUMBER: US/10/741,600
CURRENT FILING DATE: 2003-12-22
NUMBER OF SEQ ID NOS: 73997
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17651
LENGTH: 141121
TYPE: DNA
ORGANISM: Homo sapiens
US-10-741-600-17651

Query Match
Best Local Similarity 100.0%; Score 48; DB 8; Length 141121;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3075 ATTGTGCCACTGCACTCCAGCTGGGCAACAGCAAGCTGTCTC 3122
Db 8800 ATTGTGCCACTGCACTCCAGCTGGGCAACAGCAAGCTGTCTC 8847

RESULT 123
US-10-087-192-790
Sequence 790, Application US/10087192
Publication No. US20020182586A1
GENERAL INFORMATION:
APPLICANT: Morris, David W.
APPLICANT: Engelhard, Eric K.
TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
FILE REFERENCES: 529452000122
CURRENT APPLICATION NUMBER: US/10/087,192
CURRENT FILING DATE: 2002-03-01
PRIOR APPLICATION NUMBER: US 09/747,377
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 2059
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 790
LENGTH: 181684
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
```

```
NAME/KEY: misc.feature
LOCATION: (1) _.(181684)
OTHER INFORMATION: n = A,T,C or G
US-10-087-192-790

Query Match
Best Local Similarity 100.0%; Score 48; DB 5; Length 181684;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTGTCTC 3120
Db 155726 AGATTGTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTGTCTC 155773

RESULT 124
US-10-357-930-16322
Sequence 16322, Application US/10357930
Publication No. US20040259086A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Endege, Wilbon
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
FILE REFERENCES: MRI-007BCN
CURRENT APPLICATION NUMBER: US/10/357,930
CURRENT FILING DATE: 2003-02-04
PRIOR APPLICATION NUMBER: 09/785,276
PRIOR FILING DATE: 2003-02-16
PRIOR APPLICATION NUMBER: 60/183,319
PRIOR FILING DATE: 2000-02-17
PRIOR APPLICATION NUMBER: 60/189,862
PRIOR FILING DATE: 2000-03-16
PRIOR APPLICATION NUMBER: 60/207,454
PRIOR FILING DATE: 2000-05-25
PRIOR APPLICATION NUMBER: 60/211,314
PRIOR FILING DATE: 2000-06-09
PRIOR APPLICATION NUMBER: 60/219,007
PRIOR FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/255,281
PRIOR FILING DATE: 2000-12-13
NUMBER OF SEQ ID NOS: 62232
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 16322
LENGTH: 440
TYPE: DNA
ORGANISM: Homo sapiens
US-10-357-930-16322

Query Match
Best Local Similarity 100.0%; Score 47; DB 8; Length 440;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3070 GCAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTCTC 3116
Db 378 GCAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTCTC 424

RESULT 125
US-10-357-930-46148
Sequence 46148, Application US/10357930
Publication No. US20040259086A1
GENERAL INFORMATION:
APPLICANT: Schlegel, Robert
APPLICANT: Endege, Wilbon
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
FILE REFERENCES: MRI-007BCN
CURRENT APPLICATION NUMBER: US/10/357,930
CURRENT FILING DATE: 2003-02-04
```

```

; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46148
; LENGTH: 516
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 495..496
; OTHER INFORMATION: n = A,T,C or G
; US-10-357-930-46148

Query Match      1.5%; Score 47; DB 8; Length 516;
Best Local Similarity 100.0%; Pred. No. 2,2e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      3070 GCAAGATTGTGCGACTGCACCTCCAGGCTGGGCAACAGACAGACTCT 3116
DB      418 GCAAGATTGTGCGACTGCACCTCCAGGCTGGGCAACAGACAGACTCT 464

RESULT 126
US-09-925-065A-440136/C
; Sequence 440136, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 440136
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-440136
```

```

Query Match      1.5%; Score 47; DB 4; Length 534;
Best Local Similarity 100.0%; Pred. No. 2,2e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      3071 CAAAGTTTGGCGACTGCACCTCCAGGCTGGGCAACAGACAGACTCT 3117
DB      161 CAAAGTTTGGCGACTGCACCTCCAGGCTGGGCAACAGACAGACTCT 115
```

```

RESULT 127
US-09-925-065A-440137/C
; Sequence 440137, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 440137
; LENGTH: 534
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-440137

Query Match      1.5%; Score 47; DB 4; Length 534;
Best Local Similarity 100.0%; Pred. No. 2,2e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      3071 CAAAGTTTGGCGACTGCACCTCCAGGCTGGGCAACAGACAGACTCT 3117
DB      161 CAAAGTTTGGCGACTGCACCTCCAGGCTGGGCAACAGACAGACTCT 115
```

```

RESULT 128
US-09-925-065A-623360/C
; Sequence 623360, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 623360
; LENGTH: 566
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-925-065A-623360
```

```

Query Match      1.5%; Score 47; DB 4; Length 566;
Best Local Similarity 100.0%; Pred. No. 2,2e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```

QY      3076 TTGCGCACTGCACCTCCAGGCTGGGCAACAGACAGACTCTGTCTC 3122
```



```
Db          336 TTGTGCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTGTCTC 290

RESULT 129
US-10-027-632-115306/c
; Sequence 115306, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115306
; LENGTH: 3231
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-115306

Query Match          1.5%; Score 47; DB 5; Length 3231;
Best Local Similarity 100.0%; Pred. No. 2.1e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          3076 TTGTGCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTGTCTC 3122
Db          840 TTGTGCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTGTCTC 794

RESULT 130
US-10-027-632-115307/c
; Sequence 115307, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115307
```

```
; LENGTH: 3231
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-115307

Query Match          1.5%; Score 47; DB 5; Length 3231;
Best Local Similarity 100.0%; Pred. No. 2.1e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          3076 TTGTGCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTGTCTC 3122
Db          840 TTGTGCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTGTCTC 794

RESULT 131
US-10-027-632-115308/c
; Sequence 115308, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115308
; LENGTH: 3231
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-115308

Query Match          1.5%; Score 47; DB 5; Length 3231;
Best Local Similarity 100.0%; Pred. No. 2.1e-13;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy          3076 TTGTGCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTGTCTC 3122
Db          840 TTGTGCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTGTGTCTC 794

RESULT 132
US-10-027-632-115306/c
; Sequence 115306, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
```



FILE REFERENCE: ORIN-003CIP  
CURRENT APPLICATION NUMBER: US/10/674,124A  
PRIOR FILING DATE: 2003-09-26  
PRIOR APPLICATION NUMBER: 10/257,511  
PRIOR FILING DATE: 2003-03-07  
PRIOR APPLICATION NUMBER: PCT/JP00/07621  
PRIOR FILING DATE: 2000-10-30  
PRIOR APPLICATION NUMBER: JP2000-112699  
PRIOR FILING DATE: 2000-04-13  
PRIOR APPLICATION NUMBER: JP2002-327516  
PRIOR FILING DATE: 2002-09-28  
PRIOR APPLICATION NUMBER: JP2002-383869  
PRIOR FILING DATE: 2002-12-09  
NUMBER OF SEQ ID NOS: 27110  
SEQ ID NO 25317  
LENGTH: 322  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: AC007064.27\_66627  
FEATURE:  
OTHER INFORMATION: Located on chromosome 22  
FEATURE:  
OTHER INFORMATION: Distance between a terminus base of telomere on  
OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base  
OTHER INFORMATION: sequence : 14352146  
FEATURE:  
OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of  
OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and  
OTHER INFORMATION: 5'-terminus of this base sequence : 34953  
US-10-674-124A-25317

Query Match 1.5%; Score 46; DB 8; Length 322;  
Best Local Similarity 100.0%; Pred.No. 7.3e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3076 TTGTGCACCTGCACCTCGAGCTGGGCAACAGACAAAGACTGTCT 3121  
Db 153 TTGTGCACCTGCACCTCGAGCTGGGCAACAGACAAAGACTGTCT 108

RESULT 137  
US-10-914-037-282/C  
Sequence 282, Application US/10914037  
Publication No. US20050003444A1  
GENERAL INFORMATION:  
APPLICANT: Nehls, Michael  
APPLICANT: Zambrowicz, Brian  
TITLE OF INVENTION: NOVEL HUMAN POLYNUCLEOTIDES AND THE  
TITLE OF INVENTION: POLYPEPTIDES ENCODED THEREBY  
FILE REFERENCE: 8535-0029-999  
CURRENT APPLICATION NUMBER: US/10/914,037  
PRIOR FILING DATE: 2004-08-06  
PRIOR APPLICATION NUMBER: US/09/428,674  
PRIOR FILING DATE: 1999-10-27  
PRIOR APPLICATION NUMBER: US 60/106,442  
PRIOR FILING DATE: 1998-10-30  
NUMBER OF SEQ ID NOS: 1008  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 282  
LENGTH: 380  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(380)  
OTHER INFORMATION: n = A,T,C or G  
US-10-914-037-282

Query Match 1.5%; Score 46; DB 8; Length 380;  
Best Local Similarity 100.0%; Pred.No. 7.3e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCACCTGCACCTCGAGCTGGGCAACAGACAAAGACTC 3116  
Db 49 CAAGATTGTGCACCTGCACCTCGAGCTGGGCAACAGACAAAGACTC 4

RESULT 138  
US-09-867-701-6376/C  
Sequence 6376, Application US/09867701  
Patent No. US20020132237A1  
GENERAL INFORMATION:  
APPLICANT: Aglate, Paul A.  
APPLICANT: Jones, Robert  
APPLICANT: Harlocker, Susan L.  
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY  
TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER  
FILE REFERENCE: 210121.497  
CURRENT APPLICATION NUMBER: US/09/867,701  
PRIOR FILING DATE: 2001-05-29  
NUMBER OF SEQ ID NOS: 10912  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 6376  
LENGTH: 405  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-867-701-6376

Query Match 1.5%; Score 46; DB 3; Length 405;  
Best Local Similarity 100.0%; Pred.No. 7.3e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCACCTGCACCTCGAGCTGGGCAACAGACAAAGACTC 3116  
Db 106 CAAGATTGTGCACCTGCACCTCGAGCTGGGCAACAGACAAAGACTC 61

RESULT 139  
US-10-674-124A-24232/C  
Sequence 24232, Application US/10674124A  
Publication No. US20040197797A1  
GENERAL INFORMATION:  
APPLICANT: INOKO, Hidetoshi  
APPLICANT: TAMURA, Gen  
TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE  
TITLE OF INVENTION: GENETIC POLYMORPHISM MARKERS  
FILE REFERENCE: ORIN-003CIP  
CURRENT APPLICATION NUMBER: US/10/674,124A  
PRIOR FILING DATE: 2003-09-26  
PRIOR APPLICATION NUMBER: 10/257,511  
PRIOR FILING DATE: 2003-03-07  
PRIOR APPLICATION NUMBER: PCT/JP00/07621  
PRIOR FILING DATE: 2000-10-30  
PRIOR APPLICATION NUMBER: JP2000-112699  
PRIOR FILING DATE: 2000-04-13  
PRIOR APPLICATION NUMBER: JP2002-327516  
PRIOR FILING DATE: 2002-09-28  
PRIOR APPLICATION NUMBER: JP2002-383869  
PRIOR FILING DATE: 2002-12-09  
NUMBER OF SEQ ID NOS: 27110  
SEQ ID NO 24232  
LENGTH: 440  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: AC020909.4\_102225  
FEATURE:  
OTHER INFORMATION: Located on chromosome 19  
FEATURE:  
OTHER INFORMATION: Distance between a terminus base of telomere on  
OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base  
OTHER INFORMATION: sequence : 67910743  
FEATURE:  
OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of

Query Match 1.5%; Score 46; DB 3; Length 405;  
Best Local Similarity 100.0%; Pred.No. 7.3e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and  
OTHER INFORMATION: 5'-terminus of this base sequence : 99683  
US-10-674-124A-24232

Query Match 1.5%; Score 46; DB 8; Length 440;  
Best Local Similarity 100.0%; Pred. No. 7.3e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCATCTGAGGCCGAGATTGAGACCAAGCCTGGCCAACAT 2940  
|||||  
DB 335 GGTGATCATCTGAGGCCGAGATTGAGACCAAGCCTGGCCAACAT 290

## RESULT 140

US-09-814-353-17260/c  
Sequence 17260, Application US/09814353  
Publication No. US20030165831A1

GENERAL INFORMATION:  
APPLICANT: Thompson, Pamela  
APPLICANT: Little, James  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND  
TITLE OF INVENTION: THERAPY OF OVARIAN CANCER  
FILE REFERENCE: MRI-006B

CURRENT APPLICATION NUMBER: US/09/814,353  
CURRENT FILING DATE: 2001-03-21  
PRIOR APPLICATION NUMBER: US 60/191,031  
PRIOR FILING DATE: 2000-03-21  
PRIOR APPLICATION NUMBER: US 60/207,124  
PRIOR FILING DATE: 2000-05-25  
PRIOR APPLICATION NUMBER: US 60/211,940  
PRIOR FILING DATE: 2000-06-15  
PRIOR APPLICATION NUMBER: US 60/216,820  
PRIOR FILING DATE: 2000-07-07  
PRIOR APPLICATION NUMBER: US 60/220,661  
PRIOR FILING DATE: 2000-07-25  
PRIOR APPLICATION NUMBER: US 60/257,672  
PRIOR FILING DATE: 2000-12-21  
NUMBER OF SEQ ID NOS: 22037  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 17260  
LENGTH: 458  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-814-353-17260

Query Match 1.5%; Score 46; DB 3; Length 458;  
Best Local Similarity 100.0%; Pred. No. 7.3e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCATCTGAGGCCGAGATTGAGACCAAGCCTGGCCAACAT 2940  
|||||  
DB 405 GGTGATCATCTGAGGCCGAGATTGAGACCAAGCCTGGCCAACAT 360

## RESULT 141

US-10-674-124A-16717  
Sequence 16717, Application US/10674124A  
Publication No. US2004019797A1

GENERAL INFORMATION:  
APPLICANT: INOKO, Hideochi  
APPLICANT: TAMURA, Gen  
TITLE OF INVENTION: GENE MAPPING METHOD USING MICROSATELLITE  
FILE REFERENCE: ORIN-003CIP  
CURRENT APPLICATION NUMBER: US/10/674,124A  
CURRENT FILING DATE: 2003-09-26

PRIOR APPLICATION NUMBER: 10/257,511  
PRIOR FILING DATE: 2003-03-07  
PRIOR APPLICATION NUMBER: PCT/JP00/07621  
PRIOR FILING DATE: 2000-10-30  
PRIOR APPLICATION NUMBER: JP2000-112699

PRIOR FILING DATE: 2000-04-13  
PRIOR APPLICATION NUMBER: JP2002-327516  
PRIOR FILING DATE: 2002-09-28  
PRIOR APPLICATION NUMBER: JP2002-383869  
PRIOR FILING DATE: 2002-12-09  
NUMBER OF SEQ ID NOS: 27110  
SEQ ID NO 16717  
LENGTH: 482  
TYPE: DNA  
ORGANISM: Homo sapiens

FEATURE:  
OTHER INFORMATION: AC016744.5\_76819  
FEATURE:  
OTHER INFORMATION: Located on chromosome 10

FEATURE:  
OTHER INFORMATION: Distance between a terminus base of telomere on  
OTHER INFORMATION: chromosomal short arm and 5'-terminus of this base  
FEATURE:  
OTHER INFORMATION: sequence : 111450655  
OTHER INFORMATION: Distance between 3'-terminus of neighbour sequence of  
OTHER INFORMATION: sequence listing upward to telomere on chromosomal short arm and  
OTHER INFORMATION: 5'-terminus of this base sequence : 136558  
US-10-674-124A-16717

Query Match 1.5%; Score 46; DB 8; Length 482;  
Best Local Similarity 100.0%; Pred. No. 7.3e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCAGTGCAGCTCCAGGCTTGAGCAAGAGCAAGACTC 3116  
|||||  
DB 312 CAAGATTGTGCAGTGCAGCTCCAGGCTTGAGCAAGAGCAAGACTC 357

RESULT 142  
US-10-357-930-60554  
Sequence 60554, Application US/10357930  
Publication No. US20040259086A1

GENERAL INFORMATION:  
APPLICANT: Schlegel, Robert  
APPLICANT: Endege, Wilson  
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR  
TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF  
TITLE OF INVENTION: HUMAN PROSTATE CANCER  
FILE REFERENCE: MRI-007BCN  
CURRENT APPLICATION NUMBER: US/10/357,930  
CURRENT FILING DATE: 2003-02-04  
PRIOR APPLICATION NUMBER: 09/785,276  
PRIOR FILING DATE: 2003-02-16  
PRIOR APPLICATION NUMBER: 60/183,319  
PRIOR FILING DATE: 2000-02-17  
PRIOR APPLICATION NUMBER: 60/189,862  
PRIOR FILING DATE: 2000-03-16  
PRIOR APPLICATION NUMBER: 60/207,454  
PRIOR FILING DATE: 2000-05-25  
PRIOR APPLICATION NUMBER: 60/211,314  
PRIOR FILING DATE: 2000-06-09  
PRIOR APPLICATION NUMBER: 60/219,007  
PRIOR FILING DATE: 2000-07-18  
PRIOR APPLICATION NUMBER: 60/255,281  
PRIOR FILING DATE: 2000-12-13  
NUMBER OF SEQ ID NOS: 62232  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 60554  
LENGTH: 497  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-357-930-60554

Query Match 1.5%; Score 46; DB 8; Length 497;  
Best Local Similarity 100.0%; Pred. No. 7.3e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACCTGAGGCCAGAGTTCCAGACCAGCCTGGCCAACAT 2940  
|||  
Db 192 GGTGATCACCTGAGGCCAGAGTTCCAGACCAGCCTGGCCAACAT 237

RESULT 143  
US-10-027-632-91304  
; Sequence 91304, Application US/10027632  
; Publication NO. US20020198371A1  
; GENERAL INFORMATION:

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91304
;
; LENGTH: 535
;
; TYPE: DNA
;
; ORGANISM: Human
US-10-027-632-91304

```

```

RESULT 144
US-10-027-632-91305
? Sequence 91305, Application US/10027632
? Publication No. US20020198371A1
? GENERAL INFORMATION:
? APPLICANT: Wang, David G.
? TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
? TITLE OF INVENTION: Polymorphisms in the Human Genome
? FILE REFERENCE: 108827, 129
? CURRENT APPLICATION NUMBER: US/10/027,632
? CURRENT FILING DATE: 2002-04-30
? PRIOR APPLICATION NUMBER: US 60/218, 006
? PRIOR FILING DATE: 2000-07-12
? PRIOR APPLICATION NUMBER: US 60/198,676
? PRIOR FILING DATE: 2000-04-20
? PRIOR APPLICATION NUMBER: US 60/193,483
? PRIOR FILING DATE: 2000-03-29
? PRIOR APPLICATION NUMBER: US 60/185,218
? PRIOR FILING DATE: 2000-02-24
? PRIOR APPLICATION NUMBER: US 60/167,363
? PRIOR FILING DATE: 1999-11-23
? PRIOR APPLICATION NUMBER: US 60/156,358
? PRIOR FILING DATE: 1999-09-28
? PRIOR APPLICATION NUMBER: US 60/146,002
? PRIOR FILING DATE: 1999-08-09
? NUMBER OF SEQ ID NOS: 325720

```

```

; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91305
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-91305

```

Query Match	1.54;	Score 46;	DB 5;	Length 535;
Best Local Similarity	100.04;	Pred. No. 7.3e-13;		
Matches 46;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

RESULT 145  
US-10-027-632-91306  
; Sequence 91306, Application US/10027632  
; Publication No. US20020198371A1  
; Publication Information

! ORGANISM: Human  
US-10-027-632-91306

RESULT 146  
US-10-027-632-91304  
/ Sequence 91304, Application US/10027632  
/ Publication No. US20030204075A9  
/ GENERAL INFORMATION:  
/ APPLICANT: Wang, David G.  
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
/ POLYMORPHISMS IN THE HUMAN GENOME  
/ TITLE OF INVENTION: Polymorphisms in the Human Genome  
/ FILE REFERENCE: 108927.129  
/ CURRENT APPLICATION NUMBER: US/10/027.632  
/ CURRENT FILING DATE: 2002-04-30  
/ PRIOR APPLICATION NUMBER: US 60/218,006  
/ PRIOR FILING DATE: 2000-07-12  
/ PRIOR APPLICATION NUMBER: US 60/198,676  
/ PRIOR FILING DATE: 2000-04-20

```

; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91304
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-91304
```

```
Query Match          1.5%: Score 46; DB 6; Length 535;
Best Local Similarity 100.0%; Pred. No. 7.3e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2895 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAAAT 2940
DB      383 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAAAT 428
```

```
RESULT 147
US-10-027-632-91305
; Sequence 91305, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91305
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-91305
```

```
Query Match          1.5%: Score 46; DB 6; Length 535;
Best Local Similarity 100.0%; Pred. No. 7.3e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2895 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAAAT 2940
DB      383 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAAAT 428
```

```
RESULT 148
US-10-027-632-91306
; Sequence 91306, Application US/10027632
; Publication No. US20030204075A9
```

```

; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 91306
; LENGTH: 535
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-91306
```

```
Query Match          1.5%: Score 46; DB 6; Length 535;
Best Local Similarity 100.0%; Pred. No. 7.3e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2895 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAAAT 2940
DB      383 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAAAT 428
```

```
RESULT 149
US-09-925-065A-124249
; Sequence 124249, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 124249
; LENGTH: 539
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-124249
```

```
Query Match          1.5%: Score 46; DB 4; Length 539;
Best Local Similarity 100.0%; Pred. No. 7.2e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2895 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAAAT 2940
```

Db 358 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCTGGCCAACAT 403

RESULT 150  
US-09-925-065A-843811/c  
; Sequence 843811, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIORITY FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 843811  
; LENGTH: 606  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-843811

Query Match 1.5%; Score 46; DB 4; Length 606;  
Best Local Similarity 100.0%; Pred. No. 7.2e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCTGGCCAACAT 2940  
|||||  
Db 298 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCTGGCCAACAT 253

RESULT 151  
US-09-925-065A-801712/c  
; Sequence 801712, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIORITY FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 801712  
; LENGTH: 609  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-801712

Query Match 1.5%; Score 46; DB 4; Length 609;  
Best Local Similarity 100.0%; Pred. No. 7.2e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCTGGCCAACAT 2940  
|||||  
Db 299 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCTGGCCAACAT 254

RESULT 152  
US-09-925-065A-48663/c  
; Sequence 48663, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIORITY FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 48663  
; LENGTH: 610  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-48663

Query Match 1.5%; Score 46; DB 4; Length 610;  
Best Local Similarity 100.0%; Pred. No. 7.2e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCTGGCCAACAT 2940  
|||||  
Db 540 GGTGATCACCCTGAGGCCAGAGTTGAGACCAAGCTGGCCAACAT 495

RESULT 153  
US-09-925-065A-65044  
; Sequence 65044, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIORITY FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 65044  
; LENGTH: 610  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-65044



Query Match 1.5%; Score 46; DB 4; Length 610;  
Best Local Similarity 100.0%; Pred. No. 7.2e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACTGAGCCGAGAGTTGAGACCAAGCTGGCCAACTG 2940  
DB 240 GGTGATCACTGAGCCGAGAGTTGAGACCAAGCTGGCCAACTG 285

## RESULT 154

US-09-925-065A-47411/c  
; Sequence 47410, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 47410  
; LENGTH: 614  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-47410

Query Match 1.5%; Score 46; DB 4; Length 614;  
Best Local Similarity 100.0%; Pred. No. 7.2e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGACCTGAGGCAAGACAGCAAGACTCTG 3118  
DB 264 AGATTGTGCACCTGACCTGAGGCAAGACAGCAAGACTCTG 219

## RESULT 155

US-09-925-065A-47411/c  
; Sequence 47411, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 47411  
; LENGTH: 614  
; TYPE: DNA

; ORGANISM: Homo sapiens  
US-09-925-065A-47411

Query Match 1.5%; Score 46; DB 4; Length 614;  
Best Local Similarity 100.0%; Pred. No. 7.2e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGACCTGAGGCAAGACAGCAAGACTCTG 3118  
DB 264 AGATTGTGCACCTGACCTGAGGCAAGACAGCAAGACTCTG 219

## RESULT 156

US-09-925-065A-930543/c  
; Sequence 930543, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 930543  
; LENGTH: 615  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-930543

Query Match 1.5%; Score 46; DB 4; Length 615;  
Best Local Similarity 100.0%; Pred. No. 7.2e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2897 TGAATCACTGAGGCCAGAGGTTGAGACCAAGCTGGCCAACTG 2942  
DB 290 TGAATCACTGAGGCCAGAGGTTGAGACCAAGCTGGCCAACTG 245

## RESULT 157

US-09-925-065A-930544/c  
; Sequence 930544, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:

; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0

```

; SEQ ID NO 930544
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-930544

```

```

Query Match
Best Local Similarity 1.5%; Score 46; DB 4; Length 615;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 2897 TGGATCAGCTGAGGCGGAGGAGTTTCAGACCAAGCTGGCCCAACTAG 2942
DB 290 TGGATCAGCTGAGGCGGAGGAGTTTCAGACCAAGCTGGCCCAACTAG 245

```

```

RESULT 158
US-09-925-065A-870312/c
; Sequence 870312, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 870312
; LENGTH: 616
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-870312

```

```

Query Match
Best Local Similarity 1.5%; Score 46; DB 4; Length 616;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3071 CAAGATTGTGCGCACTGCACTCCAGCTGGGCAAGAGCAAGACTC 3116
DB 246 CAAGATTGTGCGCACTGCACTCCAGCTGGGCAAGAGCAAGACTC 201

```

```

RESULT 159
US-10-027-632-115128
; Sequence 115128, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363

```

```

; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115128
; LENGTH: 621
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-115128

```

```

Query Match
Best Local Similarity 1.5%; Score 46; DB 5; Length 621;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3071 CAAGATTGTGCGCACTGCACTCCAGCTGGGCAAGAGCAAGACTC 3116
DB 364 CAAGATTGTGCGCACTGCACTCCAGCTGGGCAAGAGCAAGACTC 409

```

```

RESULT 160
US-10-027-632-115129
; Sequence 115129, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 115129
; LENGTH: 621
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-115129

```

```

Query Match
Best Local Similarity 1.5%; Score 46; DB 5; Length 621;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3071 CAAGATTGTGCGCACTGCACTCCAGCTGGGCAAGAGCAAGACTC 3116
DB 364 CAAGATTGTGCGCACTGCACTCCAGCTGGGCAAGAGCAAGACTC 409

```

```

RESULT 161
US-10-027-632-115128
; Sequence 115128, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129

```

```
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 115128
/ LENGTH: 621
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-115128
```

```
Query Match
Best Local Similarity 1.5%; Score 46; DB 6; Length 621;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 3071 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGCTC 3116
Db 364 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGCTC 409
```

```
RESULT 162
US-10-027-632-115129
/ Sequence 115129, Application US/10027632
/ Publication No. US20030204075A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
/ FILE REFERENCE: 108827.129
/ CURRENT APPLICATION NUMBER: US/10/027,632
/ CURRENT FILING DATE: 2002-04-30
/ PRIOR APPLICATION NUMBER: US 60/218,006
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: US 60/198,676
/ PRIOR FILING DATE: 2000-04-20
/ PRIOR APPLICATION NUMBER: US 60/193,483
/ PRIOR FILING DATE: 2000-03-29
/ PRIOR APPLICATION NUMBER: US 60/185,218
/ PRIOR FILING DATE: 2000-02-24
/ PRIOR APPLICATION NUMBER: US 60/167,363
/ PRIOR FILING DATE: 1999-11-23
/ PRIOR APPLICATION NUMBER: US 60/156,358
/ PRIOR FILING DATE: 1999-09-28
/ PRIOR APPLICATION NUMBER: US 60/146,002
/ PRIOR FILING DATE: 1999-08-09
/ NUMBER OF SEQ ID NOS: 325720
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 115129
/ LENGTH: 621
/ TYPE: DNA
/ ORGANISM: Human
US-10-027-632-115129
```

```
Query Match
Best Local Similarity 1.5%; Score 46; DB 6; Length 621;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 3071 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGCTC 3116
Db 364 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGCTC 409
```

```
RESULT 163
US-09-925-065A-554324
/ Sequence 554324, Application US/09925065A
/ Publication No. US20050228172A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 554324
/ LENGTH: 654
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-554324
```

```
Query Match
Best Local Similarity 1.5%; Score 46; DB 4; Length 654;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
Qy 2895 GGTGATGACCTGAGCCAGAGGATTGAGACGAGCTGGCCAACT 2940
Db 154 GGTGATGACCTGAGCCAGAGGATTGAGACGAGCTGGCCAACT 199
```

```
RESULT 164
US-09-925-065A-759345/c
/ Sequence 759345, Application US/09925065A
/ Publication No. US20050228172A9
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 759345
/ LENGTH: 654
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-759345
```

```
Query Match
Best Local Similarity 1.5%; Score 46; DB 4; Length 654;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 3071 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTC 3116  
|||||  
Db 274 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTC 229

RESULT 165  
US-09-925-065A-790726  
; Sequence 790726, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 790726  
; LENGTH: 659  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-790726

Query Match 1.5%; Score 46; DB 4; Length 659;  
Best Local Similarity 100.0%; Pred. No. 7,2e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2895 GGTGATGACCTGAGGCCAGAGTTGAGACCAAGCTGCGCAACAT 2940  
|||||  
Db 180 GGTGATGACCTGAGGCCAGAGTTGAGACCAAGCTGCGCAACAT 225

RESULT 166  
US-09-925-065A-790727  
; Sequence 790727, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 790727  
; LENGTH: 659  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-790727

Query Match 1.5%; Score 46; DB 4; Length 659;

Best Local Similarity 100.0%; Pred. No. 7,2e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATGACCTGAGGCCAGAGTTGAGACCAAGCTGCGCAACAT 2940  
|||||  
Db 180 GGTGATGACCTGAGGCCAGAGTTGAGACCAAGCTGCGCAACAT 225

RESULT 167  
US-09-925-065A-771454  
; Sequence 771454, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 771454  
; LENGTH: 661  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-771454

Query Match 1.5%; Score 46; DB 4; Length 661;  
Best Local Similarity 100.0%; Pred. No. 7,2e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTC 3116  
|||||  
Db 388 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTC 433

RESULT 168  
US-09-925-065A-737520  
; Sequence 737520, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 737520  
; LENGTH: 672  
; TYPE: DNA  
; ORGANISM: Homo sapiens

```
US-09-925-065A-737520
Query Match          1.5%; Score 46; DB 4; Length 672;
Best Local Similarity 100.0%; Pred. No. 7.2e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTC 3116
DB 385 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTC 430

RESULT 169
US-09-925-065A-737227/c
; Sequence 737227, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; NUCLEOTIDE POLYMORPHISMS IN THE HUMAN GENOME
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 737227
; LENGTH: 680
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-737227

Query Match          1.5%; Score 46; DB 4; Length 680;
Best Local Similarity 100.0%; Pred. No. 7.2e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTC 3116
DB 294 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTC 249

RESULT 170
US-09-925-065A-852555
; Sequence 852555, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; NUCLEOTIDE POLYMORPHISMS IN THE HUMAN GENOME
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 852555

US-10-027-632-144717/c
; Sequence 144717, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; NUCLEOTIDE POLYMORPHISMS IN THE HUMAN GENOME
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23

US-09-925-065A-852556
Query Match          1.5%; Score 46; DB 4; Length 689;
Best Local Similarity 100.0%; Pred. No. 7.2e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTCTG 3118
DB 293 AGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTCTG 338

RESULT 171
US-09-925-065A-852556
; Sequence 852556, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; NUCLEOTIDE POLYMORPHISMS IN THE HUMAN GENOME
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 852556
; LENGTH: 689
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-852556
```

```

; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144717
; LENGTH: 756
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-144717

Query Match      1.5%; Score 46; DB 5; Length 756;
Best Local Similarity 100.0%; Pred. No. 7.2e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
Db      294 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 249

RESULT 173
US-10-027-632-144718/c
; Sequence 144718, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144718
; LENGTH: 756
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-144718

Query Match      1.5%; Score 46; DB 5; Length 756;
Best Local Similarity 100.0%; Pred. No. 7.2e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
Db      294 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 249

RESULT 174
US-10-027-632-144719/c
; Sequence 144719, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144719
; LENGTH: 756
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-144719
```

```

; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144719
; LENGTH: 756
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-144719

Query Match      1.5%; Score 46; DB 5; Length 756;
Best Local Similarity 100.0%; Pred. No. 7.2e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
Db      294 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 249

RESULT 175
US-10-027-632-144717/c
; Sequence 144717, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 144717
; LENGTH: 756
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-144717
```

```

Query Match      1.5%; Score 46; DB 6; Length 756;
Best Local Similarity 100.0%; Pred. No. 7.2e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
Db      294 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 249
```

RESULT 176  
US-10-027-632-144718/C  
; Sequence 144718, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1998-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 144718  
; LENGTH: 756  
; TYPE: DNA  
; ORGANISM: Human  
US-10-027-632-144718

Query Match 1.5%; Score 46; DB 6; Length 756;  
Best Local Similarity 100.0%; Pred. No. 7.2e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGTTGTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTCTG 3116  
DB 294 CAAGTTGTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTCTG 249

RESULT 177  
US-10-027-632-144719/C  
; Sequence 144719, Application US/10027632  
; Publication No. US20030204075A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide  
; TITLE OF INVENTION: Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.129  
; CURRENT APPLICATION NUMBER: US/10/027,632  
; PRIOR FILING DATE: 2002-04-30  
; PRIOR APPLICATION NUMBER: US 60/218,006  
; PRIOR FILING DATE: 2000-07-12  
; PRIOR APPLICATION NUMBER: US 60/198,676  
; PRIOR FILING DATE: 2000-04-20  
; PRIOR APPLICATION NUMBER: US 60/193,483  
; PRIOR FILING DATE: 2000-03-29  
; PRIOR APPLICATION NUMBER: US 60/185,218  
; PRIOR FILING DATE: 2000-02-24  
; PRIOR APPLICATION NUMBER: US 60/167,363  
; PRIOR FILING DATE: 1999-11-23  
; PRIOR APPLICATION NUMBER: US 60/156,358  
; PRIOR FILING DATE: 1999-09-28  
; PRIOR APPLICATION NUMBER: US 60/146,002  
; PRIOR FILING DATE: 1998-08-09  
; NUMBER OF SEQ ID NOS: 325720  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 144719  
; LENGTH: 756  
; TYPE: DNA

; ORGANISM: Human  
US-10-027-632-144719

Query Match 1.5%; Score 46; DB 6; Length 756;  
Best Local Similarity 100.0%; Pred. No. 7.2e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGTTGTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTCTG 3116  
DB 294 CAAGTTGTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTCTG 249

RESULT 178  
US-09-925-065A-938958/C  
; Sequence 938958, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 938958  
; LENGTH: 840  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-938958

Query Match 1.5%; Score 46; DB 4; Length 840;  
Best Local Similarity 100.0%; Pred. No. 7.2e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTCTG 3118  
DB 146 AGATTGTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTCTG 101

RESULT 179  
US-09-925-065A-938959/C  
; Sequence 938959, Application US/09925065A  
; Publication No. US20050228172A9  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0



SEQ ID NO 938959  
LENGTH: 840  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-925-065A-938959

Query Match 1.5%; Score 46; DB 4; Length 840;  
Best Local Similarity 100.0%; Pred. No. 7.2e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGAACTCTG 318  
DB 146 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGAACTCTG 101

RESULT 180  
US-09-925-065A-938960/c  
Sequence 938960, Application US/09925065A  
Publication No. US20050228172A9

GENERAL INFORMATION:  
APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single  
FILE REFERENCE: 108827.135  
CURRENT APPLICATION NUMBER: US/09/925,065A

CURRENT FILING DATE: 2001-08-08  
PRIOR APPLICATION NUMBER: US 60/243,096

PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: US 60/252,147

PRIOR FILING DATE: 2000-11-20  
PRIOR APPLICATION NUMBER: US 60/250,092

PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: US 60/261,766

PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/289,846

NUMBER OF SEQ ID NOS: 957086  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 938960  
LENGTH: 840  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-925-065A-938960

Query Match 1.5%; Score 46; DB 4; Length 840;  
Best Local Similarity 100.0%; Pred. No. 7.2e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGAACTCTG 318  
DB 146 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGAACTCTG 101

RESULT 181  
US-09-925-065A-711495  
Sequence 711495, Application US/09925065A  
Publication No. US20050228172A9

GENERAL INFORMATION:  
APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single  
FILE REFERENCE: 108827.135  
CURRENT APPLICATION NUMBER: US/09/925,065A

CURRENT FILING DATE: 2001-08-08  
PRIOR APPLICATION NUMBER: US 60/243,096

PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: US 60/252,147

PRIOR FILING DATE: 2000-11-20  
PRIOR APPLICATION NUMBER: US 60/250,092

PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: US 60/261,766  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/289,846

PRIOR FILING DATE: 2001-05-09  
NUMBER OF SEQ ID NOS: 957086  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 711495  
LENGTH: 1105  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-925-065A-711495

Query Match 1.5%; Score 46; DB 4; Length 1105;  
Best Local Similarity 100.0%; Pred. No. 7.1e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGTTGTGCACCTGCACTCCAGCTGGGCAACAGAGAACTCTG 316  
DB 874 CAAGTTGTGCACCTGCACTCCAGCTGGGCAACAGAGAACTCTG 919

RESULT 182  
US-09-925-065A-678054  
Sequence 678054, Application US/09925065A  
Publication No. US20050228172A9

GENERAL INFORMATION:  
APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single  
FILE REFERENCE: 108827.135  
CURRENT APPLICATION NUMBER: US/09/925,065A

CURRENT FILING DATE: 2001-08-08  
PRIOR APPLICATION NUMBER: US 60/243,096

PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: US 60/252,147

PRIOR FILING DATE: 2000-11-20  
PRIOR APPLICATION NUMBER: US 60/250,092

PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: US 60/261,766

PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/289,846

NUMBER OF SEQ ID NOS: 957086  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 678054  
LENGTH: 1364  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-925-065A-678054

Query Match 1.5%; Score 46; DB 4; Length 1364;  
Best Local Similarity 100.0%; Pred. No. 7.1e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGAACTCTG 318  
DB 351 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGAGAACTCTG 396

RESULT 183  
US-09-925-065A-689300  
Sequence 689300, Application US/09925065A  
Publication No. US20050228172A9

GENERAL INFORMATION:  
APPLICANT: Wang, David G.

TITLE OF INVENTION: Identification and Mapping of Single  
FILE REFERENCE: 108827.135  
CURRENT APPLICATION NUMBER: US/09/925,065A

CURRENT FILING DATE: 2001-08-08  
PRIOR APPLICATION NUMBER: US 60/243,096

PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: US 60/252,147

PRIOR FILING DATE: 2000-11-20  
PRIOR APPLICATION NUMBER: US 60/250,092

PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: US 60/289,846

```
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 689300
; LENGTH: 1635
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-689300

Query Match          1.5% Score 46; DB 4; Length 1635;
Best Local Similarity 100.0%; Pred. No. 7e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3071 CAAGATTGTGCGCACTGCAGCTCCAGCTGGGCAACAGAGCAAGACTC 3116
Db      831 CAAGATTGTGCGCACTGCAGCTCCAGCTGGGCAACAGAGCAAGACTC 876

RESULT 184
US-09-925-065A-677066
; Sequence 677066, Application US/09925065A
; Publication No. US20050228172A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: Nucleotide Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 677066
; LENGTH: 2227
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-677066

Query Match          1.5% Score 46; DB 4; Length 2227;
Best Local Similarity 100.0%; Pred. No. 7e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      2895 GGTGATACCTGAGGCGAGAGCTTCGAGACCAAGCTGGCCAAACAT 2940
Db      2035 GGTGATACCTGAGGCGAGAGCTTCGAGACCAAGCTGGCCAAACAT 2080

RESULT 185
US-10-027-632-260094
; Sequence 260094, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
```

```
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 260094
; LENGTH: 2798
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-260094

Query Match          1.5% Score 46; DB 5; Length 2798;
Best Local Similarity 100.0%; Pred. No. 7e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3071 CAAGATTGTGCGCACTGCAGCTCCAGCTGGGCAACAGAGCAAGACTC 3116
Db      1434 CAAGATTGTGCGCACTGCAGCTCCAGCTGGGCAACAGAGCAAGACTC 1479

RESULT 186
US-10-027-632-260094
; Sequence 260094, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; FILE REFERENCE: Polymorphisms in the Human Genome
; CURRENT APPLICATION NUMBER: US/10/027,632
; PRIOR FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 260094
; LENGTH: 2798
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-260094

Query Match          1.5% Score 46; DB 6; Length 2798;
Best Local Similarity 100.0%; Pred. No. 7e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3071 CAAGATTGTGCGCACTGCAGCTCCAGCTGGGCAACAGAGCAAGACTC 3116
Db      1434 CAAGATTGTGCGCACTGCAGCTCCAGCTGGGCAACAGAGCAAGACTC 1479

RESULT 187
US-10-008-789-10/c
; Sequence 10, Application US/10008789
```

```

Publication No. US20030125276A1
GENERAL INFORMATION:
APPLICANT: Kenneth Dobie
TITLE OF INVENTION: ANTISENSE MODULATION OF THYROID HORMONE RECEPTOR INTERACTOR 6 EXF
FILE REFERENCE: R1S-0333
CURRENT APPLICATION NUMBER: US/10/008,789
CURRENT FILING DATE: 2001-11-08
NUMBER OF SEQ ID NOS: 89
SEQ ID NO 10
LENGTH: 7001
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: exon
LOCATION: (486)...(740)
OTHER INFORMATION: exon 1
NAME/KEY: exon:intron junction
LOCATION: (740)...(741)
OTHER INFORMATION: exon 1:intron 1
NAME/KEY: intron
LOCATION: (741)...(994)
OTHER INFORMATION: intron 1
NAME/KEY: intron:exon junction
LOCATION: (994)...(995)
OTHER INFORMATION: intron 1:exon 2
NAME/KEY: exon
LOCATION: (995)...(1122)
OTHER INFORMATION: exon 2
NAME/KEY: exon:intron junction
LOCATION: (1122)...(1123)
OTHER INFORMATION: exon 2:intron 2
NAME/KEY: intron
LOCATION: (1123)...(1241)
OTHER INFORMATION: intron 2
NAME/KEY: intron:exon junction
LOCATION: (1241)...(1242)
OTHER INFORMATION: intron 2:exon 3
NAME/KEY: exon
LOCATION: (1242)...(1367)
OTHER INFORMATION: exon 3
NAME/KEY: exon:intron junction
LOCATION: (1367)...(1368)
OTHER INFORMATION: exon 3:intron 3
NAME/KEY: intron
LOCATION: (1368)...(1628)
OTHER INFORMATION: intron 3
NAME/KEY: intron:exon junction
LOCATION: (1628)...(1629)
OTHER INFORMATION: intron 3:exon 4
NAME/KEY: exon
LOCATION: (1629)...(2000)
OTHER INFORMATION: exon 4
NAME/KEY: exon:intron junction
LOCATION: (2000)...(2001)
OTHER INFORMATION: exon 4:intron 4
NAME/KEY: intron
LOCATION: (2001)...(3503)
OTHER INFORMATION: intron 4
NAME/KEY: intron:exon junction
LOCATION: (3503)...(3504)
OTHER INFORMATION: intron 4:exon 5
NAME/KEY: exon
LOCATION: (3504)...(3597)
OTHER INFORMATION: exon 5
NAME/KEY: exon:intron junction
LOCATION: (3597)...(3598)
OTHER INFORMATION: exon 5:intron 5
NAME/KEY: intron
LOCATION: (3598)...(3707)
OTHER INFORMATION: intron 5
NAME/KEY: intron:exon junction
LOCATION: (3707)...(3708)

```

```

OTHER INFORMATION: intron 5:exon 6
NAME/KEY: exon
LOCATION: (3708)...(3877)
OTHER INFORMATION: exon 6
NAME/KEY: exon:intron junction
LOCATION: (3877)...(3878)
OTHER INFORMATION: exon 6:intron 6
NAME/KEY: intron
LOCATION: (3878)...(4676)
OTHER INFORMATION: intron 6
NAME/KEY: intron:exon junction
LOCATION: (4676)...(4677)
OTHER INFORMATION: intron 6:exon 7
NAME/KEY: exon
LOCATION: (4677)...(4855)
OTHER INFORMATION: exon 7
NAME/KEY: exon:intron junction
LOCATION: (4855)...(4856)
OTHER INFORMATION: exon 7:intron 7
NAME/KEY: intron
LOCATION: (4856)...(5757)
OTHER INFORMATION: intron 7
NAME/KEY: intron:exon junction
LOCATION: (5757)...(5758)
OTHER INFORMATION: intron 7:exon 8
NAME/KEY: exon
LOCATION: (5758)...(5878)
OTHER INFORMATION: exon 8
NAME/KEY: exon:intron junction
LOCATION: (5878)...(5879)
OTHER INFORMATION: exon 8:intron 8
NAME/KEY: intron
LOCATION: (5879)...(6305)
OTHER INFORMATION: intron 8
NAME/KEY: intron:exon junction
LOCATION: (6305)...(6306)
OTHER INFORMATION: intron 8:exon 9
NAME/KEY: exon
LOCATION: (6306)...(6650)
OTHER INFORMATION: exon 9
US-10-008-789-10

Query Match 1.5%; Score 46; DB 6; Length 7001;
Best Local Similarity 100.0%; Pred. No. 6.8e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCACCTGCATCCAGCTGGGCAACAGCAAGACTC 3116
DB 6054 CAAGATTGTGCACCTGCATCCAGCTGGGCAACAGCAAGACTC 6009

RESULT 188
US-11-071-724-10/c
Sequence 10, Application US/11071724
Publication No. US20050222073A1
GENERAL INFORMATION:
APPLICANT: C. Frank Bennett
APPLICANT: Brett P. Monia
APPLICANT: Rosanne Crooke
APPLICANT: Mark Graham
APPLICANT: Susan M. Freiler
APPLICANT: Pamela Nero
APPLICANT: Edward Mancewicz
APPLICANT: Nicholas M. Dean
APPLICANT: Lex M. Cowsett
APPLICANT: Kenneth Dobie
TITLE OF INVENTION: ANTISENSE MODULATION OF THYROID HORMONE RECEPTOR INTERACTOR 6 EXF
FILE REFERENCE: BNDL-0018US.P1
CURRENT APPLICATION NUMBER: US/11/071,724
CURRENT FILING DATE: 2005-03-02
PRIOR APPLICATION NUMBER: 10/190,366
PRIOR FILING DATE: 2002-07-02
PRIOR APPLICATION NUMBER: 10/008,789

```

```
/ PRIOR FILING DATE: 2001-11-08
/ PRIOR APPLICATION NUMBER: 11/031,827
/ PRIOR FILING DATE: 2005-01-07
/ PRIOR APPLICATION NUMBER: 09/925,139
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: 10/114,544
/ PRIOR FILING DATE: 2002-04-01
/ PRIOR APPLICATION NUMBER: 10/019,470
/ PRIOR FILING DATE: 2002-05-09
/ PRIOR APPLICATION NUMBER: PCT/US00/19019
/ PRIOR FILING DATE: 2000-07-12
/ PRIOR APPLICATION NUMBER: 09/357,071
/ PRIOR FILING DATE: 1999-07-19
/ NUMBER OF SEQ ID NOS: 89
/ SEQ ID NO 10
/ LENGTH: 7001
/ TYPE: DNA
/ ORGANISM: Homo sapiens
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: (486)...(740)
/ OTHER INFORMATION: exon 1
/ FEATURE:
/ NAME/KEY: exon:intron junction
/ LOCATION: (740)...(741)
/ OTHER INFORMATION: exon 1:intron 1
/ FEATURE:
/ NAME/KEY: intron
/ LOCATION: (741)...(994)
/ OTHER INFORMATION: intron 1
/ FEATURE:
/ NAME/KEY: intron:exon junction
/ LOCATION: (994)...(995)
/ OTHER INFORMATION: intron 1:exon 2
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: (995)...(1122)
/ OTHER INFORMATION: exon 2
/ FEATURE:
/ NAME/KEY: exon:intron junction
/ LOCATION: (1122)...(1123)
/ OTHER INFORMATION: exon 2:intron 2
/ FEATURE:
/ NAME/KEY: intron
/ LOCATION: (1123)...(1241)
/ OTHER INFORMATION: intron 2
/ FEATURE:
/ NAME/KEY: intron:exon junction
/ LOCATION: (1241)...(1242)
/ OTHER INFORMATION: intron 2:exon 3
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: (1242)...(1367)
/ OTHER INFORMATION: exon 3
/ FEATURE:
/ NAME/KEY: exon:intron junction
/ LOCATION: (1367)...(1368)
/ OTHER INFORMATION: exon 3:intron 3
/ FEATURE:
/ NAME/KEY: intron
/ LOCATION: (1368)...(1628)
/ OTHER INFORMATION: intron 3
/ FEATURE:
/ NAME/KEY: intron:exon junction
/ LOCATION: (1628)...(1629)
/ OTHER INFORMATION: intron 3:exon 4
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: (1629)...(2000)
/ OTHER INFORMATION: exon 4
/ FEATURE:
/ NAME/KEY: exon:intron junction
/ LOCATION: (2000)...(2001)
/ OTHER INFORMATION: exon 4:intron 4
/ FEATURE:
/ NAME/KEY: intron
/ LOCATION: (2001)...(3503)
/ OTHER INFORMATION: intron 4
/ FEATURE:
/ NAME/KEY: intron:exon junction
/ LOCATION: (3503)...(3504)
/ OTHER INFORMATION: intron 4:exon 5
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: (3504)...(3597)
/ OTHER INFORMATION: exon 5
/ FEATURE:
/ NAME/KEY: exon:intron junction
/ LOCATION: (3597)...(3598)
/ OTHER INFORMATION: exon 5:intron 5
/ FEATURE:
/ NAME/KEY: intron
/ LOCATION: (3598)...(3707)
/ OTHER INFORMATION: intron 5
/ FEATURE:
/ NAME/KEY: intron:exon junction
/ LOCATION: (3707)...(3708)
/ OTHER INFORMATION: intron 5:exon 6
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: (3708)...(3877)
/ OTHER INFORMATION: exon 6
/ FEATURE:
/ NAME/KEY: exon:intron junction
/ LOCATION: (3877)...(3878)
/ OTHER INFORMATION: exon 6:intron 6
/ FEATURE:
/ NAME/KEY: intron
/ LOCATION: (3878)...(4676)
/ OTHER INFORMATION: intron 6
/ FEATURE:
/ NAME/KEY: intron:exon junction
/ LOCATION: (4676)...(4677)
/ OTHER INFORMATION: intron 6:exon 7
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: (4677)...(4855)
/ OTHER INFORMATION: exon 7
/ FEATURE:
/ NAME/KEY: exon:intron junction
/ LOCATION: (4855)...(4856)
/ OTHER INFORMATION: exon 7:intron 7
/ FEATURE:
/ NAME/KEY: intron
/ LOCATION: (4856)...(5757)
/ OTHER INFORMATION: intron 7
/ FEATURE:
/ NAME/KEY: intron:exon junction
/ LOCATION: (5757)...(5758)
/ OTHER INFORMATION: intron 7:exon 8
/ FEATURE:
/ NAME/KEY: exon
/ LOCATION: (5758)...(5878)
/ OTHER INFORMATION: exon 8
/ FEATURE:
/ NAME/KEY: exon:intron junction
/ LOCATION: (5878)...(5879)
/ OTHER INFORMATION: exon 8:intron 8
/ FEATURE:
/ NAME/KEY: intron
/ LOCATION: (5879)...(6305)
/ OTHER INFORMATION: intron 8
/ FEATURE:
/ NAME/KEY: intron:exon junction
/ LOCATION: (6305)...(6306)
/ OTHER INFORMATION: intron 8:exon 9
```

FEATURE:  
NAME/KEY: exon  
LOCATION: (6306)...(6650)  
OTHER INFORMATION: exon 9  
US-11-071-724-10

Query Match 1.5%; Score 46; DB 10; Length 7001;  
Best Local Similarity 100.0%; Pred. No. 6.8e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTC 3116  
DB 6054 CAAGATTGTGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTC 6009

RESULT 189  
US-09-764-891-9601  
Sequence 9601, Application US/09764891  
Publication No. US2003007808A1

GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC006  
CURRENT APPLICATION NUMBER: US/09/764,891  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 10231  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 9601  
LENGTH: 13409  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-764-891-9601

Query Match 1.5%; Score 46; DB 3; Length 13409;  
Best Local Similarity 100.0%; Pred. No. 6.7e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTC 3116  
DB 12121 CAAGATTGTGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTC 12166

RESULT 190  
US-10-205-428-930  
Sequence 930, Application US/10205428  
Publication No. US20030108907A1

GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PA117C1  
CURRENT APPLICATION NUMBER: US/10/205,428  
CURRENT FILING DATE: 2002-07-26  
PRIOR APPLICATION NUMBER: 09/764,892  
PRIOR FILING DATE: 2001-01-17  
PRIOR APPLICATION NUMBER: 60/179,065  
PRIOR FILING DATE: 2000-01-31  
PRIOR APPLICATION NUMBER: 60/180,628  
PRIOR FILING DATE: 2000-02-04  
PRIOR APPLICATION NUMBER: 60/214,886  
PRIOR FILING DATE: 2000-06-28  
PRIOR APPLICATION NUMBER: 60/217,487  
PRIOR FILING DATE: 2000-07-11  
PRIOR APPLICATION NUMBER: 60/225,758  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/220,963  
PRIOR FILING DATE: 2000-07-26  
PRIOR APPLICATION NUMBER: 60/217,496  
PRIOR FILING DATE: 2000-07-11  
PRIOR APPLICATION NUMBER: 60/225,447  
PRIOR FILING DATE: 2000-08-14  
PRIOR APPLICATION NUMBER: 60/218,290  
PRIOR FILING DATE: 2000-07-14

Remaining Prior Application data removed - See File Wrapper or PALM.  
NUMBER OF SEQ ID NOS: 1019  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 930  
LENGTH: 13409  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-205-428-930

Query Match 1.5%; Score 46; DB 5; Length 13409;  
Best Local Similarity 100.0%; Pred. No. 6.7e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTC 3116  
DB 12121 CAAGATTGTGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTC 12166

RESULT 191  
US-09-764-847-1916  
Sequence 1916, Application US/09764847  
Patent No. US20020132767A1

GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC009  
CURRENT APPLICATION NUMBER: US/09/764,847  
CURRENT FILING DATE: 2001-01-17  
Prior application data removed - consult PALM or file wrapper  
NUMBER OF SEQ ID NOS: 2003  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1916  
LENGTH: 18501  
TYPE: DNA  
ORGANISM: Homo sapiens

FEATURE:  
NAME/KEY: SITE  
LOCATION: (9022)  
OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: SITE  
LOCATION: (9023)  
OTHER INFORMATION: n equals a,t,g, or c  
US-09-764-847-1916

Query Match 1.5%; Score 46; DB 3; Length 18501;  
Best Local Similarity 100.0%; Pred. No. 6.6e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTC 3116  
DB 224 CAAGATTGTGCCACTGCACCTCCAGCTGGGCAACAGAGCAAGACTC 269

RESULT 192  
US-10-092-154-1916  
Sequence 1916, Application US/10092154  
Publication No. US20030054375A1

GENERAL INFORMATION:  
APPLICANT: Rosen et al.  
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
FILE REFERENCE: PC009C1  
CURRENT APPLICATION NUMBER: US/10/092,154  
CURRENT FILING DATE: 2002-03-07  
NUMBER OF SEQ ID NOS: 2003  
Prior application removed - See File Wrapper or Palm  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 1916  
LENGTH: 18501  
TYPE: DNA  
ORGANISM: Homo sapiens

FEATURE:  
NAME/KEY: misc feature  
LOCATION: (9022)

OTHER INFORMATION: n equals a,t,g, or c  
NAME/KEY: misc\_feature  
LOCATION: (3023)  
OTHER INFORMATION: n equals a,t,g, or c  
US-10-092-154-1916

Query Match 1.5%; Score 46; DB 5; Length 18501;  
Best Local Similarity 100.0%; Pred. No. 6.6e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116  
Db 224 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 2659

RESULT 193  
US-10-187-659A-11/c  
Sequence 11, Application US/10187659A  
Publication No. US20040002152A1  
GENERAL INFORMATION:  
APPLICANT: Kenneth W. Doble  
TITLE OF INVENTION: ANTISENSE MODULATION OF P2X4 EXPRESSION  
FILE REFERENCE: RTS-0379  
CURRENT APPLICATION NUMBER: US/10/187,659A  
CURRENT FILING DATE: 2002-07-01  
NUMBER OF SEQ ID NOS: 143  
SEQ ID NO 11  
LENGTH: 25001  
TYPE: DNA  
ORGANISM: H. sapiens  
US-10-187-659A-11

Query Match 1.5%; Score 46; DB 6; Length 25001;  
Best Local Similarity 100.0%; Pred. No. 6.6e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2895 GGTGATGACCTGAGGCGCAGAGTTCGAGACCAAGCCTGGCCAACAT 2940  
Db 2633 GGTGATGACCTGAGGCGCAGAGTTCGAGACCAAGCCTGGCCAACAT 2588

RESULT 194  
US-10-455-552-1/c  
Sequence 1, Application US/10455552  
Publication No. US20040018533A1  
GENERAL INFORMATION:  
APPLICANT: Adam, Gail Isabel  
APPLICANT: Landown, Maria  
APPLICANT: Roth, Richard  
APPLICANT: Denisenko, Mikhail  
TITLE OF INVENTION: DIAGNOSING PREDISPOSITION TO FAT  
TITLE OF INVENTION: DEPOSITION AND THERAPEUTIC METHODS FOR REDUCING FAT  
TITLE OF INVENTION: DEPOSITION AND TREATMENT OF ASSOCIATED CONDITIONS  
FILE REFERENCE: 52459-20030.00  
CURRENT APPLICATION NUMBER: US/10/455,552  
CURRENT FILING DATE: 2003-06-04  
PRIOR APPLICATION NUMBER: US 60/386,012  
PRIOR FILING DATE: 2002-06-04  
NUMBER OF SEQ ID NOS: 98  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1  
LENGTH: 28616  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-455-552-1

Query Match 1.5%; Score 46; DB 7; Length 28616;  
Best Local Similarity 100.0%; Pred. No. 6.6e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2895 GGTGATGACCTGAGGCGCAGAGTTCGAGACCAAGCCTGGCCAACAT 2940  
|||||

Db 4783 GGTGATGACCTGAGGCCAGAGATTCTGAGACCAAGCCTGGCCAACAT 4738

RESULT 195  
US-10-719-993-7030/c  
Sequence 7030, Application US/10719993  
Publication No. US20040265849A1  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
TITLE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001496  
CURRENT APPLICATION NUMBER: US/10/719,993  
CURRENT FILING DATE: 2003-11-24  
NUMBER OF SEQ ID NOS: 55342  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 7030  
LENGTH: 57095  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-719-993-7030

Query Match 1.5%; Score 46; DB 8; Length 57095;  
Best Local Similarity 100.0%; Pred. No. 6.5e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116  
Db 23272 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 23227

RESULT 196  
US-10-322-281-526/c  
Sequence 526, Application US/10322281  
Publication No. US20040126762A1  
GENERAL INFORMATION:  
APPLICANT: David W. Morris  
APPLICANT: Marc S. Malandro  
TITLE OF INVENTION: Novel Compositions and Methods in Cancer  
FILE REFERENCE: 529452001000  
CURRENT APPLICATION NUMBER: US/10/322,281  
CURRENT FILING DATE: 2002-12-17  
NUMBER OF SEQ ID NOS: 866  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 526  
LENGTH: 58922  
TYPE: DNA  
ORGANISM: Homo sapiens  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: (1)...(58922)  
OTHER INFORMATION: n = A,T,C or G  
US-10-322-281-526

Query Match 1.5%; Score 46; DB 7; Length 58922;  
Best Local Similarity 100.0%; Pred. No. 6.4e-13;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116  
Db 53826 CAAGATTGTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 53781

RESULT 197  
US-10-741-601-5619/c  
Sequence 5619, Application US/10741601  
Publication No. US20040166519A1  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
TITLE OF INVENTION: STENOSIS, METHODS OF DETECTION AND USES THEREOF  
FILE REFERENCE: CL001500  
CURRENT APPLICATION NUMBER: US/10/741,601

```

; CURRENT FILING DATE: 2003-12-22
; NUMBER OF SEQ ID NOS: 26415
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 5619
; LENGTH: 59914
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(59914)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-741-601-5619

Query Match          1.5%; Score 46; DB 7; Length 59914;
Best Local Similarity 100.0%; Pred. No. 6.4e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2895 GGTGATCAGCTGAGGCCAGAGATTGAGACCAAGCCCTGGCCAAACAT 2940
DB      10387 GGTGATCAGCTGAGGCCAGAGATTGAGACCAAGCCCTGGCCAAACAT 10342

RESULT 198
US-10-719-993-6871/c
; Sequence 6871, Application US/10719993
; Publication No. US20040265849A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; FILE OF INVENTION: ALZHEIMER'S DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001496
; CURRENT APPLICATION NUMBER: US/10/719,993
; CURRENT FILING DATE: 2003-11-24
; NUMBER OF SEQ ID NOS: 55342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 6871
; LENGTH: 93011
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(93011)
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-
US-10-719-993-6871

Query Match          1.5%; Score 46; DB 8; Length 93011;
Best Local Similarity 100.0%; Pred. No. 6.4e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2895 GGTGATCAGCTGAGGCCAGAGATTGAGACCAAGCCCTGGCCAAACAT 2940
DB      82157 GGTGATCAGCTGAGGCCAGAGATTGAGACCAAGCCCTGGCCAAACAT 82112

RESULT 199
US-10-052-482-232
; Sequence 232, Application US/10052482
; Publication No. US20040072264A1
; GENERAL INFORMATION:
; APPLICANT: Engelhard, Eric
; APPLICANT: Morris, David
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: A-71087/RMS/DCF
; CURRENT APPLICATION NUMBER: US/10/052,482
; CURRENT FILING DATE: 2002-08-15
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 241
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 232
; LENGTH: 96595
```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (49183)..(49621)
; OTHER INFORMATION: "n" at positions 49183 to 49621 can be any base
US-10-052-482-232

Query Match          1.5%; Score 46; DB 7; Length 96595;
Best Local Similarity 100.0%; Pred. No. 6.4e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3071 CAAGATTGGCCACTGCACTCCAGCTGGGCAACAGAGACAGACTC 3116
DB      57090 CAAGATTGGCCACTGCACTCCAGCTGGGCAACAGAGACAGACTC 57135

RESULT 200
US-10-704-513-1/c
; Sequence 1, Application US/10704513
; Publication No. US20050170500A1
; GENERAL INFORMATION:
; APPLICANT: ROTH, RICHARD B.
; APPLICANT: NELSON, MATTHEW ROBERTS
; APPLICANT: KAMMERER, STEFAN M.
; APPLICANT: BRAUN, ANDREAS
; TITLE OF INVENTION: METHODS FOR IDENTIFYING RISK OF MELANOMA AND TREATMENTS
; FILE OF INVENTION: THEROF
; FILE REFERENCE: SEQ-4062-OT
; CURRENT APPLICATION NUMBER: US/10/704,513
; CURRENT FILING DATE: 2003-11-06
; PRIOR APPLICATION NUMBER: 60/489,703
; PRIOR FILING DATE: 2003-07-23
; PRIOR APPLICATION NUMBER: 60/424,475
; PRIOR FILING DATE: 2002-11-06
; NUMBER OF SEQ ID NOS: 774
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 1
; LENGTH: 99100
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (47431)
; OTHER INFORMATION: a, c, g, t, unknown, or other
US-10-704-513-1

Query Match          1.5%; Score 46; DB 9; Length 99100;
Best Local Similarity 100.0%; Pred. No. 6.4e-13;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3071 CAAGATTGGCCACTGCACTCCAGCTGGGCAACAGAGACAGACTC 3116
DB      16521 CAAGATTGGCCACTGCACTCCAGCTGGGCAACAGAGACAGACTC 16476
```

Search completed: May 11, 2006, 06:40:25  
JOB time : 2402 secs



GenCore version 5.1.8  
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 11, 2006, 05:54:19 ; Search time 1343 Seconds  
(without alignments)  
9481.115 Million cell updates/sec

Title: US-10-760-320A-102  
Perfect score: 3122  
Sequence: 1 actagagctgggggttagcgc.....acagagaacactctgtctc 3122

Scoring table: OLIGO\_NWC  
Gapop 60.0 , Gapext 60.0

Searched: 9312410 seqs, 2039259788 residues

Word size : 1

Total number of hits satisfying chosen parameters: 18623998

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 500 summaries

Database : Published Applications NA.New.\*

1: /SIDSS/prodata/1/pubpna/US08\_NEW\_PUB.seq1.\*  
2: /SIDSS/prodata/1/pubpna/US06\_NEW\_PUB.seq2.\*  
3: /SIDSS/prodata/1/pubpna/US07\_NEW\_PUB.seq3.\*  
4: /SIDSS/prodata/1/pubpna/US08\_NEW\_PUB.seq4.\*  
5: /SIDSS/prodata/1/pubpna/PCR\_NEW\_PUB.seq5.\*  
6: /SIDSS/prodata/1/pubpna/US09\_NEW\_PUB.seq1.\*  
7: /SIDSS/prodata/1/pubpna/US09\_NEW\_PUB.seq2.\*  
8: /SIDSS/prodata/1/pubpna/US09\_NEW\_PUB.seq3.\*  
9: /SIDSS/prodata/1/pubpna/US10\_NEW\_PUB.seq1.\*  
10: /SIDSS/prodata/1/pubpna/US10\_NEW\_PUB.seq2.\*  
11: /SIDSS/prodata/1/pubpna/US10\_NEW\_PUB.seq3.\*  
12: /SIDSS/prodata/1/pubpna/US10\_NEW\_PUB.seq4.\*  
13: /SIDSS/prodata/1/pubpna/US10\_NEW\_PUB.seq5.\*  
14: /SIDSS/prodata/1/pubpna/US11\_NEW\_PUB.seq1.\*  
15: /SIDSS/prodata/1/pubpna/US11\_NEW\_PUB.seq2.\*  
16: /SIDSS/prodata/1/pubpna/US11\_NEW\_PUB.seq3.\*  
17: /SIDSS/prodata/1/pubpna/US11\_NEW\_PUB.seq4.\*  
18: /SIDSS/prodata/1/pubpna/US11\_NEW\_PUB.seq5.\*  
19: /SIDSS/prodata/1/pubpna/US60\_NEW\_PUB.seq1.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	509	16.3	560	7	US-09-925-065A-740956, Sequence 740956,
C 2	506	16.2	557	7	US-09-925-065A-769508, Sequence 769508,
C 3	505	16.2	556	7	US-09-925-065A-736351, Sequence 736351,
C 4	504	16.1	555	7	US-09-925-065A-737120, Sequence 737120,
C 5	53	1.7	477	11	US-10-301-480-20199, Sequence 20199, A
C 6	53	1.7	477	12	US-10-301-480-633608, Sequence 633608, A
C 7	53	1.7	492	11	US-10-301-480-20198, Sequence 20198, A
C 8	53	1.7	492	12	US-10-301-480-633607, Sequence 633607, A
C 9	52	1.7	559	7	US-09-925-065A-591947, Sequence 591947,
C 10	52	1.7	574	7	US-09-925-065A-591947, Sequence 591947,
C 11	52	1.7	592	7	US-09-925-065A-59682, Sequence 59682, A
C 12	52	1.7	592	11	US-10-301-480-160920, Sequence 160920,
C 13	52	1.7	592	12	US-10-301-480-774329, Sequence 774329,
C 14	52	1.7	614	12	US-10-301-480-454942, Sequence 454942,

C 15	52	1.7	614	12	US-10-301-480-1068351, Sequence 1068351,
C 16	52	1.7	617	7	US-09-925-065A-386361, Sequence 386361,
C 17	52	1.7	618	12	US-10-301-480-611724, Sequence 611724,
C 18	52	1.7	618	12	US-10-301-480-1225133, Sequence 1225133,
C 19	52	1.7	695	7	US-09-925-065A-708701, Sequence 708701,
C 20	52	1.7	974	12	US-10-301-480-598972, Sequence 598972,
C 21	52	1.7	974	12	US-10-301-480-1212381, Sequence 1212381,
C 22	52	1.7	18996	17	US-11-121-086-44, Sequence 84, Appl
C 23	52	1.7	137671	17	US-11-121-086-47, Sequence 47, Appl
C 24	52	1.7	398287	10	US-10-995-551-13396, Sequence 13396, A
C 25	51	1.6	385	7	US-09-925-065A-111249, Sequence 111249,
C 26	51	1.6	561	7	US-09-925-065A-141833, Sequence 141833,
C 27	51	1.6	561	12	US-10-301-480-236672, Sequence 236672,
C 28	51	1.6	561	12	US-10-301-480-555081, Sequence 850081,
C 29	51	1.6	587	11	US-10-301-480-1011, Sequence 1011, Ap
C 30	51	1.6	587	11	US-10-301-480-42180, Sequence 42180, A
C 31	51	1.6	587	12	US-10-301-480-614420, Sequence 614420,
C 32	51	1.6	587	12	US-10-301-480-655589, Sequence 655589,
C 33	51	1.6	852	12	US-10-301-480-589593, Sequence 589593,
C 34	51	1.6	852	12	US-10-301-480-1203002, Sequence 1203002,
C 35	51	1.6	197096	17	US-11-121-086-107, Sequence 107, App
C 36	50	1.6	571	7	US-09-925-065A-484883, Sequence 484883,
C 37	50	1.6	572	7	US-09-925-065A-785367, Sequence 785367,
C 38	50	1.6	625	7	US-09-925-065A-894777, Sequence 894777,
C 39	50	1.6	630	7	US-09-925-065A-892664, Sequence 892664,
C 40	50	1.6	633	7	US-09-925-065A-892405, Sequence 892405,
C 41	50	1.6	633	7	US-09-925-065A-892514, Sequence 892514,
C 42	50	1.6	633	7	US-09-925-065A-949927, Sequence 949927,
C 43	50	1.6	676	7	US-09-925-065A-917225, Sequence 917225,
C 44	50	1.6	676	7	US-09-925-065A-917226, Sequence 917226,
C 45	50	1.6	676	7	US-09-925-065A-944569, Sequence 944569,
C 46	50	1.6	739	7	US-09-925-065A-934608, Sequence 934608,
C 47	50	1.6	742	7	US-09-925-065A-941367, Sequence 941367,
C 48	50	1.6	999	12	US-10-301-480-567391, Sequence 567391,
C 49	50	1.6	999	12	US-10-301-480-1180800, Sequence 1180800,
C 50	50	1.6	1986	11	US-10-301-480-100096, Sequence 100096,
C 51	50	1.6	1986	12	US-10-301-480-113505, Sequence 113505,
C 52	50	1.6	187745	17	US-11-121-086-83, Sequence 83, Appl
C 53	50	1.6	207908	17	US-11-121-086-83, Sequence 83, Appl
C 54	50	1.6	212805	17	US-11-121-086-83, Sequence 83, Appl
C 55	49	1.6	417	12	US-10-301-480-835645, Sequence 835645,
C 56	49	1.6	417	12	US-10-301-480-849054, Sequence 849054,
C 57	49	1.6	420	7	US-09-925-065A-140670, Sequence 140670,
C 58	49	1.6	530	11	US-10-301-480-15264, Sequence 15264, A
C 59	49	1.6	530	12	US-10-301-480-628673, Sequence 628673,
C 60	49	1.6	530	12	US-10-301-480-849055, Sequence 849055,
C 61	49	1.6	530	12	US-10-301-480-849055, Sequence 849055,
C 62	49	1.6	532	7	US-09-925-065A-140671, Sequence 140671,
C 63	49	1.6	552	7	US-09-925-065A-18609, Sequence 18609, A
C 64	49	1.6	552	7	US-09-925-065A-18610, Sequence 18610, A
C 65	49	1.6	552	7	US-09-925-065A-18611, Sequence 18611, A
C 66	49	1.6	552	11	US-10-301-480-119846, Sequence 119846,
C 67	49	1.6	552	11	US-10-301-480-119847, Sequence 119847,
C 68	49	1.6	552	11	US-10-301-480-119848, Sequence 119848,
C 69	49	1.6	552	12	US-10-301-480-733255, Sequence 733255,
C 70	49	1.6	552	12	US-10-301-480-733256, Sequence 733256,
C 71	49	1.6	552	12	US-10-301-480-733257, Sequence 733257,
C 72	49	1.6	695	7	US-09-925-065A-708700, Sequence 708700,
C 73	49	1.6	750	7	US-09-925-065A-932036, Sequence 932036,
C 74	49	1.6	750	7	US-09-925-065A-932036, Sequence 932036,
C 75	49	1.6	243328	13	US-10-508-307-14, Sequence 14, Appl
C 76	48	1.5	437	11	US-10-301-480-43305, Sequence 24305, A
C 77	48	1.5	437	11	US-10-301-480-637714, Sequence 637714, A
C 78	48	1.5	438	11	US-10-301-480-433148, Sequence 433148, A
C 79	48	1.5	438	12	US-10-301-480-656557, Sequence 656557,
C 80	48	1.5	617	12	US-10-301-480-597620, Sequence 597620,
C 81	48	1.5	617	12	US-10-301-480-1211029, Sequence 1211029,
C 82	48	1.5	802	12	US-10-301-480-597621, Sequence 597621,
C 83	48	1.5	802	12	US-10-301-480-1211030, Sequence 1211030,
C 84	48	1.5	19875	17	US-11-124-368A-2878, Sequence 2878, Ap
C 85	47	1.5	534	7	US-09-925-065A-440136, Sequence 440136,
C 86	47	1.5	534	7	US-09-925-065A-440137, Sequence 440137,
C 87	47	1.5	566	7	US-09-925-065A-623360, Sequence 623360,

88	46	1.5	201	10	US-10-995-561-64399	Sequence 64399, A	161	45	1.4	568	12	US-10-301-480-411598	Sequence 411598,
C 89	46	1.5	496	11	US-10-301-480-20200	Sequence 20200, A	162	45	1.4	568	12	US-10-301-480-727744,	Sequence 727744,
C 90	46	1.5	496	12	US-10-301-480-633609	Sequence 633609,	163	45	1.4	568	12	US-10-301-480-1025007	Sequence 1025007,
C 91	46	1.5	514	12	US-10-301-480-570118	Sequence 570118,	C 164	45	1.4	590	7	US-09-925-065A-371064	Sequence 371064,
C 92	46	1.5	514	12	US-10-301-480-1183527	Sequence 1183527,	165	45	1.4	592	7	US-09-925-065A-927764	Sequence 927764,
C 93	46	1.5	540	12	US-10-301-480-57517	Sequence 577517,	166	45	1.4	592	7	US-09-925-065A-927765	Sequence 927765,
C 94	46	1.5	540	12	US-10-301-480-1190926	Sequence 1190926,	167	45	1.4	592	7	US-09-925-065A-949474	Sequence 949474,
C 95	46	1.5	559	7	US-09-925-065A-124292	Sequence 124292,	C 168	45	1.4	596	7	US-09-925-065A-695847	Sequence 695847,
C 96	46	1.5	559	7	US-10-301-480-321987	Sequence 321987,	C 169	45	1.4	598	7	US-09-925-065A-10882	Sequence 10882, A
C 97	46	1.5	568	12	US-10-301-480-835386	Sequence 835386,	C 170	45	1.4	598	11	US-10-301-480-112129	Sequence 112129,
C 98	46	1.5	606	7	US-09-925-065A-843811	Sequence 843811,	C 171	45	1.4	598	12	US-10-301-480-725538	Sequence 725538,
C 99	46	1.5	609	7	US-09-925-065A-801712	Sequence 801712,	C 172	45	1.4	603	12	US-10-301-480-253035	Sequence 253035,
C 100	46	1.5	610	7	US-09-925-065A-48663	Sequence 48663, A	C 173	45	1.4	603	12	US-10-301-480-866444	Sequence 866444,
C 101	46	1.5	610	7	US-09-925-065A-65044	Sequence 65044, A	C 174	45	1.4	604	7	US-09-925-065A-160227	Sequence 160227,
C 102	46	1.5	610	11	US-10-301-480-149901	Sequence 149901,	C 175	45	1.4	604	7	US-09-925-065A-799096	Sequence 799096,
C 103	46	1.5	610	11	US-10-301-480-166282	Sequence 166282,	C 176	45	1.4	605	7	US-09-925-065A-795229	Sequence 795229,
C 104	46	1.5	610	12	US-10-301-480-763310	Sequence 763310,	C 177	45	1.4	605	7	US-09-925-065A-850785	Sequence 850785,
C 105	46	1.5	610	12	US-10-301-480-779691	Sequence 779691,	C 178	45	1.4	606	12	US-10-301-480-441075	Sequence 441075,
C 106	46	1.5	614	7	US-09-925-065A-47410	Sequence 47410, A	C 179	45	1.4	606	12	US-10-301-480-105484	Sequence 105484,
C 107	46	1.5	614	7	US-09-925-065A-47411	Sequence 47411, A	C 180	45	1.4	624	7	US-09-925-065A-846491	Sequence 846491,
C 108	46	1.5	614	11	US-10-301-480-148648	Sequence 148648,	C 181	45	1.4	627	7	US-09-925-065A-839499	Sequence 839499,
C 109	46	1.5	614	11	US-10-301-480-148649	Sequence 148649,	C 182	45	1.4	662	7	US-09-925-065A-675355	Sequence 675355,
C 110	46	1.5	614	12	US-10-301-480-762057	Sequence 762057	C 183	45	1.4	662	7	US-09-925-065A-675356	Sequence 675356,
C 111	46	1.5	615	12	US-10-301-480-762058	Sequence 762058,	C 184	45	1.4	713	11	US-10-301-480-34229	Sequence 34229, A
C 112	46	1.5	615	46	US-09-925-065A-930543	Sequence 930543,	C 185	45	1.4	713	12	US-10-301-480-647638	Sequence 647638,
C 113	46	1.5	615	7	US-09-925-065A-930544	Sequence 930544,	C 186	45	1.4	723	12	US-10-301-480-593973	Sequence 593973,
C 114	46	1.5	616	7	US-09-925-065A-870312	Sequence 870312,	C 187	45	1.4	723	12	US-10-301-480-1207382	Sequence 1207382,
C 115	46	1.5	654	7	US-09-925-065A-554324	Sequence 554324,	C 188	45	1.4	745	12	US-10-301-480-572662	Sequence 572662,
C 116	46	1.5	654	7	US-09-925-065A-759345	Sequence 759345,	C 189	45	1.4	745	12	US-10-301-480-1185971	Sequence 1185971,
C 117	46	1.5	659	7	US-09-925-065A-790726	Sequence 790726,	C 190	45	1.4	793	7	US-09-925-065A-77568	Sequence 77568, A
C 118	46	1.5	659	7	US-09-925-065A-771454	Sequence 771454,	C 191	45	1.4	793	11	US-10-301-480-114761	Sequence 114761,
C 119	46	1.5	661	7	US-09-925-065A-737520	Sequence 737520,	C 192	45	1.4	793	11	US-10-301-480-178627	Sequence 178627,
C 120	46	1.5	672	7	US-09-925-065A-737227	Sequence 737227,	C 193	45	1.4	793	12	US-10-301-480-728170	Sequence 728170,
C 121	46	1.5	680	7	US-09-925-065A-852555	Sequence 852555,	C 194	45	1.4	793	12	US-10-301-480-792236	Sequence 792236,
C 122	46	1.5	689	7	US-09-925-065A-852556	Sequence 852556,	C 195	45	1.4	803	12	US-10-301-480-556311	Sequence 556311,
C 123	46	1.5	840	7	US-09-925-065A-938958	Sequence 938958,	C 196	45	1.4	803	12	US-10-301-480-1169920	Sequence 1169920,
C 124	46	1.5	840	7	US-09-925-065A-938959	Sequence 938959,	C 197	45	1.4	818	12	US-10-301-480-556318	Sequence 556318,
C 125	46	1.5	840	7	US-09-925-065A-938960	Sequence 938960,	C 198	45	1.4	818	12	US-10-301-480-1169727	Sequence 1169727,
C 126	46	1.5	889	12	US-10-301-480-601877	Sequence 601877,	C 199	45	1.4	896	7	US-09-925-065A-98887	Sequence 98887, A
C 127	46	1.5	989	12	US-10-301-480-1215286	Sequence 1215286,	C 200	45	1.4	896	7	US-10-301-480-190128	Sequence 190128,
C 128	46	1.5	989	12	US-10-301-480-612422	Sequence 612422,	C 201	45	1.4	896	11	US-10-301-480-803337	Sequence 803337,
C 129	46	1.5	989	12	US-10-301-480-1225831	Sequence 1225831,	C 202	45	1.4	896	12	US-10-301-480-582545	Sequence 582545,
C 130	46	1.5	105	7	US-09-925-065A-711495	Sequence 711495,	C 203	45	1.4	981	12	US-10-301-480-1195954	Sequence 1195954,
C 131	46	1.5	1105	7	US-09-925-065A-678054	Sequence 678054,	C 204	45	1.4	981	12	US-10-301-480-551651	Sequence 551651,
C 132	46	1.5	1364	7	US-09-925-065A-678054	Sequence 678054,	C 205	45	1.4	994	12	US-10-301-480-600766	Sequence 600766,
C 133	46	1.5	1535	7	US-09-925-065A-689300	Sequence 689300,	C 206	45	1.4	994	12	US-10-301-480-600775	Sequence 600775,
C 134	46	1.5	2227	7	US-09-925-065A-677066	Sequence 677066,	C 207	45	1.4	994	12	US-10-301-480-1214175	Sequence 1214175,
C 135	46	1.5	2798	11	US-10-301-480-94903	Sequence 94903, A	C 208	45	1.4	994	12	US-10-301-480-535489	Sequence 535489,
C 136	46	1.5	56448	10	US-10-301-480-708312	Sequence 708312,	C 209	45	1.4	997	12	US-10-301-480-551651	Sequence 551651,
C 137	46	1.5	56448	10	US-10-995-561-13369	Sequence 13369, A	C 210	45	1.4	997	12	US-10-301-480-1148898	Sequence 1148898,
C 138	46	1.5	98345	17	US-11-112-908-36	Sequence 36, Appl	C 211	45	1.4	997	12	US-10-301-480-1148898	Sequence 1148898,
C 139	46	1.5	109661	11	US-10-330-773-795	Sequence 795, Appl	C 212	45	1.4	997	12	US-10-301-480-1165060	Sequence 1165060,
C 140	46	1.5	150450	17	US-11-112-908-54	Sequence 54, Appl	C 213	45	1.4	1000	12	US-10-301-480-608404	Sequence 608404,
C 141	46	1.5	191343	17	US-11-112-908-53	Sequence 53, Appl	C 214	45	1.4	1000	12	US-10-301-480-1221813	Sequence 1221813,
C 142	45	1.4	201	10	US-10-995-561-58923	Sequence 58923, A	C 215	45	1.4	1027	7	US-09-925-065A-85535	Sequence 85535, A
C 143	45	1.4	201	10	US-10-995-561-58959	Sequence 58959, A	C 216	45	1.4	1027	7	US-09-925-065A-85536	Sequence 85536, A
C 144	45	1.4	201	10	US-10-995-561-68537	Sequence 68537, A	C 217	45	1.4	1027	11	US-10-301-480-186775	Sequence 186775,
C 145	45	1.4	524	11	US-10-301-480-49738	Sequence 49738, A	C 218	45	1.4	1027	11	US-10-301-480-186776	Sequence 186776,
C 146	45	1.4	524	11	US-10-301-480-49739	Sequence 49739, A	C 219	45	1.4	1027	12	US-10-301-480-800184	Sequence 800184,
C 147	45	1.4	524	12	US-10-301-480-663147	Sequence 663147,	C 220	45	1.4	1027	12	US-10-301-480-800185	Sequence 800185,
C 148	45	1.4	524	12	US-10-301-480-663148	Sequence 663148,	C 221	45	1.4	1233	11	US-10-301-480-31319	Sequence 31319, A
C 149	45	1.4	527	11	US-10-301-480-9996	Sequence 9996, Ap	C 222	45	1.4	1233	11	US-10-301-480-644727	Sequence 644727,
C 150	45	1.4	527	12	US-10-301-480-622505	Sequence 622505,	C 223	45	1.4	1233	12	US-10-301-480-554428	Sequence 554428,
C 151	45	1.4	545	7	US-09-925-065A-765886	Sequence 765886,	C 224	45	1.4	1233	12	US-10-301-480-644728	Sequence 644728,
C 152	45	1.4	546	7	US-09-925-065A-799095	Sequence 799095,	C 225	45	1.4	1302	7	US-09-925-065A-553464	Sequence 553464,
C 153	45	1.4	549	7	US-09-925-065A-615308	Sequence 615308,	C 226	45	1.4	1302	11	US-10-301-480-12413	Sequence 3413, A
C 154	45	1.4	559	7	US-09-925-065A-591948	Sequence 591948,	C 227	45	1.4	1302	12	US-10-301-480-645822	Sequence 645822,
C 155	45	1.4	563	7	US-09-925-065A-12438	Sequence 12438, A	C 228	45	1.4	1302	11	US-10-301-480-645823	Sequence 645823,
C 156	45	1.4	563	11	US-10-301-480-113675	Sequence 113675,	C 229	45	1.4	1305	7	US-09-925-065A-14751	Sequence 34751, A
C 157	45	1.4	563	12	US-10-301-480-727084	Sequence 727084,	C 230	45	1.4	1305	11	US-10-301-480-135889	Sequence 135889,
C 158	45	1.4	568	7	US-09-925-065A-13098	Sequence 13098, A	C 231	45	1.4	1305	11	US-10-301-480-149398	Sequence 749398,
C 159	45	1.4	568	7	US-09-925-065A-338452	Sequence 338452,	C 232	45	1.4	1305	12	US-10-301-480-149398	Sequence 749398,
C 160	45	1.4	568	11	US-10-301-480-114335	Sequence 114335,	C 233	45	1.4	1498	7	US-09-925-065A-71613	Sequence 71613, A

234	45	1.4	1498	11	US-10-301-480-172852	Sequence 172852,	C 307	44	1.4	593	7	US-09-925-065A-237003	Sequence 237003,
235	45	1.4	1498	12	US-10-301-480-786261	Sequence 786261,	C 308	44	1.4	593	7	US-09-925-065A-337005	Sequence 337005,
236	45	1.4	1792	7	US-09-925-065A-549910	Sequence 549910,	C 309	44	1.4	593	7	US-09-925-065A-685343	Sequence 685343,
237	45	1.4	1792	12	US-10-301-480-527860	Sequence 527860,	C 310	44	1.4	599	7	US-09-925-065A-902936	Sequence 902936,
238	45	1.4	1792	12	US-10-301-480-114169	Sequence 114169,	C 311	44	1.4	612	7	US-09-925-065A-119712	Sequence 119712, A
239	45	1.4	1869	7	US-09-925-065A-689963	Sequence 689963,	C 312	44	1.4	612	11	US-10-301-480-123309	Sequence 123209,
240	45	1.4	28693	10	US-10-995-561-13341	Sequence 13341, A	C 313	44	1.4	612	12	US-10-301-480-735402	Sequence 735618,
241	45	1.4	28693	10	US-10-995-561-13323	Sequence 13323, A	C 314	44	1.4	614	11	US-10-301-480-954608	Sequence 954608, A
242	45	1.4	160326	17	US-11-121-086-29	Sequence 29, App1	C 315	44	1.4	614	12	US-10-301-480-708811	Sequence 708811,
243	45	1.4	246360	17	US-11-121-086-8	Sequence 8, App1	C 316	44	1.4	616	12	US-10-301-480-534937	Sequence 534937,
244	45	1.4	285300	10	US-10-857-780-6	Sequence 6, App1	C 317	44	1.4	616	12	US-10-301-480-1148346	Sequence 1148346,
245	45	1.4	1080000	10	US-10-928-446A-1	Sequence 1, App1	C 318	44	1.4	618	7	US-09-925-065A-697941	Sequence 697942,
246	45	1.4	1080000	10	US-10-928-446A-181	Sequence 181, App	C 319	44	1.4	618	7	US-09-925-065A-697942	Sequence 697942,
247	45	1.4	1080000	10	US-10-928-446A-183	Sequence 183, App	C 320	44	1.4	618	7	US-09-925-065A-697943	Sequence 697943,
248	45	1.4	1080000	10	US-10-928-446A-185	Sequence 185, App	C 321	44	1.4	619	7	US-09-925-065A-789262	Sequence 789262,
249	45	1.4	1080000	10	US-10-928-446A-187	Sequence 187, App	C 322	44	1.4	619	7	US-09-925-065A-789263	Sequence 789263,
250	45	1.4	1080000	10	US-10-928-446A-189	Sequence 189, App	C 323	44	1.4	619	7	US-09-925-065A-789264	Sequence 789264,
251	45	1.4	1080000	10	US-10-928-446A-191	Sequence 191, App	C 324	44	1.4	619	7	US-09-925-065A-847245	Sequence 847245,
252	45	1.4	1080000	10	US-10-928-446A-193	Sequence 193, App	C 325	44	1.4	619	7	US-09-925-065A-951338	Sequence 951338,
253	45	1.4	1080000	10	US-10-928-446A-195	Sequence 195, App	C 326	44	1.4	623	7	US-09-925-065A-521713	Sequence 521713,
254	45	1.4	1080000	10	US-10-928-446A-197	Sequence 197, App	C 327	44	1.4	627	7	US-09-925-065A-545865	Sequence 545865,
255	45	1.4	1080000	10	US-10-928-446A-199	Sequence 199, App	C 328	44	1.4	627	7	US-09-925-065A-545866	Sequence 545866,
256	45	1.4	1080000	10	US-10-928-446A-201	Sequence 201, App	C 329	44	1.4	627	7	US-09-925-065A-545867	Sequence 545867,
257	45	1.4	1691140	17	US-11-091-018-1	Sequence 1, App1	C 330	44	1.4	627	7	US-09-925-065A-821309	Sequence 821309,
258	44	1.4	201	10	US-10-995-561-36364	Sequence 36364, A	C 331	44	1.4	629	7	US-10-301-480-46185	Sequence 46185, A
259	44	1.4	201	10	US-10-995-561-42346	Sequence 42346, A	C 332	44	1.4	634	11	US-10-301-480-594691	Sequence 594691,
260	44	1.4	201	17	US-11-124-367A-21173	Sequence 21173, A	C 333	44	1.4	634	12	US-10-301-480-659594	Sequence 659594,
261	44	1.4	201	17	US-11-124-367A-32954	Sequence 32954, A	C 334	44	1.4	634	12	US-10-301-480-659594	Sequence 659594,
262	44	1.4	423	7	US-09-925-065A-566681	Sequence 566681,	C 335	44	1.4	634	12	US-10-301-480-1208300	Sequence 1208300,
263	44	1.4	423	7	US-09-925-065A-566682	Sequence 566682,	C 336	44	1.4	643	7	US-09-925-065A-692017	Sequence 692017,
264	44	1.4	448	11	US-10-301-480-15039	Sequence 15039, A	C 337	44	1.4	659	7	US-09-925-065A-741517	Sequence 741517,
265	44	1.4	448	12	US-10-301-480-628448	Sequence 628448,	C 338	44	1.4	659	7	US-09-925-065A-741518	Sequence 741518,
266	44	1.4	464	7	US-09-925-065A-602339	Sequence 602339,	C 339	44	1.4	668	7	US-09-925-065A-127024	Sequence 127024,
267	44	1.4	470	7	US-09-925-065A-767971	Sequence 767971,	C 340	44	1.4	672	12	US-10-301-480-824194	Sequence 824194,
268	44	1.4	496	7	US-09-925-065A-876041	Sequence 876041,	C 341	44	1.4	672	12	US-10-301-480-824194	Sequence 824194,
269	44	1.4	496	7	US-09-925-065A-876042	Sequence 876042,	C 342	44	1.4	673	7	US-09-925-065A-823538	Sequence 823538, A
270	44	1.4	496	7	US-09-925-065A-876042	Sequence 876042,	C 343	44	1.4	673	11	US-10-301-480-133776	Sequence 133776,
271	44	1.4	496	7	US-09-925-065A-905372	Sequence 905372,	C 344	44	1.4	673	12	US-10-301-480-870785	Sequence 870785,
272	44	1.4	525	12	US-10-301-480-508206	Sequence 508206,	C 345	44	1.4	698	11	US-10-301-480-870786	Sequence 870786, A
273	44	1.4	525	12	US-10-301-480-508207	Sequence 508207,	C 346	44	1.4	698	12	US-10-301-480-70505	Sequence 700505,
274	44	1.4	525	12	US-10-301-480-1121615	Sequence 1121615,	C 347	44	1.4	716	12	US-10-301-480-609596	Sequence 609596,
275	44	1.4	525	12	US-10-301-480-1121616	Sequence 1121616,	C 348	44	1.4	716	12	US-10-301-480-1223005	Sequence 1223005,
276	44	1.4	530	7	US-09-925-065A-297221	Sequence 297221,	C 349	44	1.4	717	12	US-10-301-480-564215	Sequence 564215,
277	44	1.4	530	7	US-09-925-065A-297222	Sequence 297222,	C 350	44	1.4	717	12	US-10-301-480-117624	Sequence 117624,
278	44	1.4	537	7	US-09-925-065A-448581	Sequence 448581,	C 351	44	1.4	718	12	US-10-301-480-582616	Sequence 582616,
279	44	1.4	537	7	US-09-925-065A-448582	Sequence 448582,	C 352	44	1.4	718	12	US-10-301-480-5176025	Sequence 5176025,
280	44	1.4	539	12	US-10-301-480-611857	Sequence 611857,	C 353	44	1.4	761	12	US-10-301-480-1188043	Sequence 1188043,
281	44	1.4	539	12	US-10-301-480-1225266	Sequence 1225266,	C 354	44	1.4	761	12	US-10-301-480-1188043	Sequence 1188043,
282	44	1.4	541	12	US-10-301-480-374012	Sequence 374012,	C 355	44	1.4	783	12	US-10-301-480-582777	Sequence 582777,
283	44	1.4	541	12	US-10-301-480-987471	Sequence 987471,	C 356	44	1.4	783	12	US-10-301-480-1196186	Sequence 1196186,
284	44	1.4	556	7	US-09-925-065A-109914	Sequence 109914,	C 357	44	1.4	784	7	US-09-925-065A-327249	Sequence 927249,
285	44	1.4	557	7	US-09-925-065A-870384	Sequence 870384,	C 358	44	1.4	822	12	US-10-301-480-574635	Sequence 574635,
286	44	1.4	557	7	US-09-925-065A-870385	Sequence 870385,	C 359	44	1.4	822	12	US-10-301-480-1188044	Sequence 1188044,
287	44	1.4	557	12	US-10-301-480-597680	Sequence 597680,	C 360	44	1.4	873	11	US-10-301-480-89331	Sequence 89331, A
288	44	1.4	557	12	US-10-301-480-1211089	Sequence 1211089,	C 361	44	1.4	873	12	US-10-301-480-702740	Sequence 702740,
289	44	1.4	562	7	US-09-925-065A-78244	Sequence 78244, A	C 362	44	1.4	934	12	US-10-301-480-535654	Sequence 535654,
290	44	1.4	562	11	US-10-301-480-179483	Sequence 179483,	C 363	44	1.4	934	12	US-10-301-480-1145063	Sequence 1145063,
291	44	1.4	562	12	US-10-301-480-792892	Sequence 792892	C 364	44	1.4	949	12	US-10-301-480-535653	Sequence 535653,
292	44	1.4	563	11	US-10-301-480-15038	Sequence 15038, A	C 365	44	1.4	949	12	US-10-301-480-535652	Sequence 535652,
293	44	1.4	563	11	US-10-301-480-535653	Sequence 535653,	C 366	44	1.4	955	12	US-10-301-480-5171291	Sequence 5171291,
294	44	1.4	566	12	US-10-301-480-628447	Sequence 628447	C 367	44	1.4	955	12	US-10-301-480-1184700	Sequence 1184700,
295	44	1.4	566	12	US-10-301-480-309509	Sequence 309509,	C 368	44	1.4	971	12	US-10-301-480-603969	Sequence 603969,
296	44	1.4	577	7	US-09-925-065A-707681	Sequence 707681,	C 369	44	1.4	971	12	US-10-301-480-603969	Sequence 603969,
297	44	1.4	577	7	US-09-925-065A-707682	Sequence 707682,	C 370	44	1.4	972	12	US-10-301-480-543105	Sequence 543105,
298	44	1.4	584	12	US-10-301-480-319476	Sequence 319476,	C 371	44	1.4	972	12	US-10-301-480-543642	Sequence 543642,
299	44	1.4	584	12	US-10-301-480-319476	Sequence 319476,	C 372	44	1.4	972	12	US-10-301-480-543643	Sequence 543643,
300	44	1.4	584	12	US-10-301-480-932883	Sequence 932883,	C 373	44	1.4	972	12	US-10-301-480-1155051	Sequence 1155051,
301	44	1.4	584	12	US-10-301-480-932885	Sequence 932885,	C 374	44	1.4	972	12	US-10-301-480-1155051	Sequence 1155051,
302	44	1.4	585	12	US-10-301-480-611357	Sequence 611357,	C 375	44	1.4	972	12	US-10-301-480-1157052	Sequence 1157052,
303	44	1.4	585	12	US-10-301-480-1224766	Sequence 1224766,	C 376	44	1.4	979	12	US-10-301-480-587102	Sequence 587102,
304	44	1.4	588	7	US-09-925-065A-729173	Sequence 729173,	C 377	44	1.4	979	12	US-10-301-480-1205011	Sequence 1205011,
305	44	1.4	589	7	US-09-925-065A-914465	Sequence 914465,	C 378	44	1.4	980	12	US-10-301-480-606095	Sequence 606095,
306	44	1.4	591	7	US-09-925-065A-936386	Sequence 936386,	C 379	44	1.4	980	12	US-10-301-480-1220314	Sequence 1220314,

C 380	44	1.4	985	12	US-10-301-480-588232	Sequence 588232,	453	44	1.4	150468	17	US-11-112-908-56	Sequence 56, Appl
C 381	44	1.4	985	12	US-10-301-480-1201641	Sequence 1201641,	C 454	44	1.4	155440	13	US-10-960-414-12	Sequence 13, Appl
C 382	44	1.4	986	12	US-10-301-480-576981	Sequence 576981,	C 455	44	1.4	168656	17	US-11-112-908-59	Sequence 59, Appl
C 383	44	1.4	996	12	US-10-301-480-1190390	Sequence 1190390,	C 456	44	1.4	168753	17	US-11-181-234-1	Sequence 1, Appl
C 384	44	1.4	998	12	US-10-301-480-543884	Sequence 543884,	C 457	44	1.4	170285	17	US-11-112-908-58	Sequence 58, Appl
C 385	44	1.4	998	12	US-10-301-480-1157293	Sequence 1157293,	C 458	44	1.4	171162	17	US-11-112-908-58	Sequence 38, Appl
C 386	44	1.4	999	12	US-10-301-480-543138	Sequence 543138,	C 459	44	1.4	176602	17	US-11-131-086-25	Sequence 28, Appl
C 387	44	1.4	999	12	US-10-301-480-563356	Sequence 563356,	C 460	44	1.4	188056	17	US-11-120-925-1	Sequence 1, Appl
C 388	44	1.4	999	12	US-10-301-480-1156547	Sequence 1156547,	C 461	44	1.4	193789	17	US-11-112-908-55	Sequence 55, Appl
C 389	44	1.4	999	12	US-10-301-480-1176545	Sequence 1176545,	C 462	44	1.4	199130	10	US-10-995-561-1323	Sequence 1233, A
C 390	44	1.4	1000	12	US-10-301-480-585793	Sequence 585793,	C 463	44	1.4	220895	10	US-10-775-169-88	Sequence 88, Appl
C 391	44	1.4	1000	12	US-10-301-480-1199142	Sequence 1199142,	C 464	44	1.4	222094	10	US-10-995-561-1324	Sequence 1244, A
C 392	44	1.4	1005	7	US-09-925-065A-726992	Sequence 726992,	C 465	44	1.4	1125000	10	US-10-995-561-13286	Sequence 13286, A
C 393	44	1.4	1005	7	US-09-925-065A-726992	Sequence 726992,	C 466	44	1.4	201	10	US-10-995-561-8854	Sequence 28594, A
C 394	44	1.4	1068	7	US-09-925-065A-716304	Sequence 716304,	C 467	44	1.4	201	10	US-10-995-561-42347	Sequence 42347, A
C 395	44	1.4	1068	7	US-09-925-065A-716304	Sequence 716304,	C 468	44	1.4	201	10	US-10-995-561-19369	Sequence 49669, A
C 396	44	1.4	1068	7	US-09-925-065A-716306	Sequence 716306,	C 469	44	1.4	201	10	US-10-995-561-50808	Sequence 50808, A
C 397	44	1.4	1068	7	US-09-925-065A-716307	Sequence 716307,	C 470	44	1.4	201	10	US-10-995-561-61087	Sequence 61087, A
C 398	44	1.4	1068	7	US-09-925-065A-716308	Sequence 716308,	C 471	44	1.4	201	10	US-10-995-561-68769	Sequence 68769, A
C 399	44	1.4	1103	11	US-10-301-480-92758	Sequence 92758, A	C 472	44	1.4	201	10	US-10-995-561-80611	Sequence 80611, A
C 400	44	1.4	1103	11	US-10-301-480-92757	Sequence 92757, A	C 473	44	1.4	201	10	US-10-995-561-80638	Sequence 80638, A
C 401	44	1.4	1103	12	US-10-301-480-706167	Sequence 706167,	C 474	44	1.4	201	17	US-11-124-367A-32355	Sequence 32355, A
C 402	44	1.4	1103	12	US-10-301-480-706168	Sequence 706168,	C 475	44	1.4	444	7	US-09-925-065A-932412	Sequence 932412,
C 403	44	1.4	1122	7	US-09-925-065A-551840	Sequence 551840,	C 476	44	1.4	444	7	US-09-925-065A-951567	Sequence 951567,
C 404	44	1.4	1122	12	US-10-301-480-529790	Sequence 529790,	C 477	44	1.4	444	7	US-09-925-065A-951568	Sequence 951568,
C 405	44	1.4	1122	12	US-10-301-480-1143199	Sequence 1143199,	C 478	44	1.4	462	11	US-10-301-480-5149	Sequence 5149, Ap
C 406	44	1.4	1141	11	US-10-301-480-40080	Sequence 40080, A	C 479	44	1.4	462	12	US-10-301-480-61858	Sequence 61858,
C 407	44	1.4	1141	12	US-10-301-480-553489	Sequence 553489, A	C 480	44	1.4	463	7	US-09-925-065A-501502	Sequence 501502,
C 408	44	1.4	1228	11	US-10-301-480-44420	Sequence 44420, A	C 481	44	1.4	476	7	US-09-925-065A-815432	Sequence 815432,
C 409	44	1.4	1228	12	US-10-301-480-647829	Sequence 647829, A	C 482	44	1.4	482	7	US-09-925-065A-168216	Sequence 168216,
C 410	44	1.4	1240	11	US-10-301-480-98516	Sequence 98516, A	C 483	44	1.4	528	7	US-09-925-065A-491876	Sequence 491876,
C 411	44	1.4	1240	12	US-10-301-480-119195	Sequence 119195,	C 484	44	1.4	528	7	US-09-925-065A-91877	Sequence 91877,
C 412	44	1.4	1242	7	US-09-925-065A-712705	Sequence 712705,	C 485	44	1.4	532	7	US-09-925-065A-485618	Sequence 485618,
C 413	44	1.4	1330	7	US-09-925-065A-58322	Sequence 58322, A	C 486	44	1.4	542	7	US-09-925-065A-411190	Sequence 411190,
C 414	44	1.4	1330	7	US-09-925-065A-58323	Sequence 58323, A	C 487	44	1.4	553	7	US-09-925-065A-408993	Sequence 408993,
C 415	44	1.4	1330	7	US-09-925-065A-58324	Sequence 58324, A	C 488	44	1.4	559	7	US-09-925-065A-651202	Sequence 651202,
C 416	44	1.4	1330	11	US-10-301-480-159560	Sequence 159560,	C 489	44	1.4	567	7	US-09-925-065A-581250	Sequence 581250,
C 417	44	1.4	1330	11	US-10-301-480-159561	Sequence 159561,	C 490	44	1.4	570	7	US-09-925-065A-586025	Sequence 586025,
C 418	44	1.4	1330	11	US-10-301-480-159562	Sequence 159562,	C 491	44	1.4	572	7	US-09-925-065A-11001	Sequence 11001, A
C 419	44	1.4	1330	12	US-10-301-480-772969	Sequence 772969,	C 492	44	1.4	572	7	US-09-925-065A-516598	Sequence 516598,
C 420	44	1.4	1330	12	US-10-301-480-772970	Sequence 772970,	C 493	44	1.4	572	11	US-10-301-480-112538	Sequence 112538,
C 421	44	1.4	1330	12	US-10-301-480-772971	Sequence 772971,	C 494	44	1.4	577	11	US-10-301-480-45127	Sequence 45127, A
C 422	44	1.4	1390	7	US-09-925-065A-47642	Sequence 47642, A	C 495	44	1.4	577	12	US-10-301-480-658536	Sequence 658536,
C 423	44	1.4	1390	11	US-10-301-480-148880	Sequence 148880,	C 496	44	1.4	581	7	US-09-925-065A-762418	Sequence 762418,
C 424	44	1.4	1390	12	US-10-301-480-762289	Sequence 762289,	C 497	44	1.4	590	7	US-09-925-065A-762419	Sequence 762419,
C 425	44	1.4	1427	7	US-09-925-065A-77993	Sequence 77993, A	C 498	44	1.4	590	7	US-09-925-065A-762419	Sequence 762419,
C 426	44	1.4	1427	11	US-10-301-480-179232	Sequence 179232,	C 499	44	1.4	590	7	US-09-925-065A-762419	Sequence 762419,
C 427	44	1.4	1427	12	US-10-301-480-792641	Sequence 792641,	C 500	44	1.4	593	7	US-09-925-065A-411191	Sequence 411191,
C 428	44	1.4	1753	7	US-09-925-065A-714611	Sequence 714611,							
C 429	44	1.4	1753	7	US-09-925-065A-714612	Sequence 714612,							
C 430	44	1.4	1753	7	US-09-925-065A-714613	Sequence 714613,							
C 431	44	1.4	1753	7	US-09-925-065A-714614	Sequence 714614,							
C 432	44	1.4	1900	11	US-10-301-480-97935	Sequence 97935, A							
C 433	44	1.4	1900	12	US-10-301-480-711344	Sequence 711344,							
C 434	44	1.4	1970	7	US-09-925-065A-721388	Sequence 721388,							
C 435	44	1.4	2078	7	US-09-925-065A-16796	Sequence 16796, A							
C 436	44	1.4	2078	11	US-10-301-480-118033	Sequence 118033,							
C 437	44	1.4	2078	12	US-10-301-480-751442	Sequence 751442,							
C 438	44	1.4	2184	7	US-09-925-065A-715885	Sequence 715885,							
C 439	44	1.4	2184	7	US-09-925-065A-715886	Sequence 715886,							
C 440	44	1.4	2184	7	US-09-925-065A-715887	Sequence 715887,							
C 441	44	1.4	2184	7	US-09-925-065A-715888	Sequence 715888,							
C 442	44	1.4	2255	7	US-09-925-065A-696477	Sequence 696477,							
C 443	44	1.4	3488	7	US-09-925-065A-84868	Sequence 84868, A							
C 444	44	1.4	3488	11	US-10-301-480-186108	Sequence 186108,							
C 445	44	1.4	3488	12	US-10-301-480-799517	Sequence 799517,							
C 446	44	1.4	23943	11	US-10-330-773-919	Sequence 919, App							
C 447	44	1.4	86361	17	US-10-995-561-13364	Sequence 13364, A							
C 448	44	1.4	86361	10	US-10-995-561-13364	Sequence 13364, A							
C 449	44	1.4	88116	10	US-10-995-561-13351	Sequence 13351, A							
C 450	44	1.4	110847	17	US-11-121-086-11	Sequence 11, Appl							
C 451	44	1.4	138941	9	US-10-489-730-10	GENERAL INFORMATION							
C 452	44	1.4	139573	12	US-10-506-513-3	Sequence 3, Appl1							

## ALIGNMENTS

RESULT 1  
US-09-925-065A-740956/c  
; Sequence 740956, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925, 065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243, 096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252, 147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250, 092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261, 766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289, 846

/ PRIOR FILING DATE: 2001-05-09  
/ NUMBER OF SEQ ID NOS: 957086  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 740956  
/ LENGTH: 560  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
US-09-925-065A-740956

Query Match 16.3%; Score 509; DB 7; Length 560;  
Best Local Similarity 99.8%; Pred. No. 9,1e-70;  
Matches 559; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 199 CATCAGGCAAAATCTCCGCGGAGCTGCGCCCGCTTTTACCTGAGGCTCAAGTTTCCC 258  
DB 560 CATCAGGCAAAATCTCCGCGGAGCTGCGCCCGCTTTTACCTGAGGCTCAAGTTTCCC 501  
QY 259 CATCGTAAATAGAACTGGGTTGATCTCCGAGCGCTAACTTCAGAACTCGATGGG 318  
DB 500 CATCGTAAATAGAACTGGGTTGATCTCCGAGCGCTAACTTCAGAACTCGATGGG 441  
QY 319 GCGAAGGGAGGAGATGGGCTCACTCAAGTGACTTCCCGCGTGAAGCCCGCTTA 378  
DB 440 GCGAAGGGAGGAGATGGGCTCACTCAAGTGACTTCCCGCGTGAAGCCCGCTTA 381  
QY 379 CCACTGATCCAGGGGGTGGCAGCTCCGCGGAGACGAGGGGGTGGGGGCTTAAAGAA 438  
DB 380 CCACTGATCCAGGGGGTGGCAGCTCCGCGGAGACGAGGGGGTGGGGGCTTAAAGAA 321  
QY 439 ACCCTAACCAGCGCCCTTGGAGCGCTTAAGCGGAGCGCGGCTCTGACGCTTCTT 498  
DB 320 ACCCTAACCAGCGCCCTTGGAGCGCTTAAGCGGAGCGCGGCTCTGACGCTTCTT 261  
QY 499 GCCCGGAGTTGGCACTCCGAGAGATGGGAGCCGACCTTCAAGTTCCAGAGAGCCAC 558  
DB 260 GCCCGGAGTTGGCACTCCGAGAGATGGGAGCCGACCTTCAAGTTCCAGAGAGCCAC 201  
QY 559 CGTGAAGGCGAGGCGGTCAGAGACAGACGTTGACTCGAGTGGCGGCTTGGAGAGAT 618  
DB 200 CGTGAAGGCGAGGCGGTCAGAGACAGACGTTGACTCGAGTGGCGGCTTGGAGAGAT 141  
QY 619 GAGCAGGAGGAGCGGAGACCGCTTAACGAGGCTCCCTTGGCGCGCCCGCTCCGAGAGCG 678  
DB 140 GAGCAGGAGGAGCGGAGACCGCTTAACGAGGCTCCCTTGGCGCGCCCGCTCCGAGAGCG 81  
QY 679 CACGTCGAGGTCCTCCGCGGCGGCTCCGTCGAGCTTGGCGGTCGCGCCGACGATCAGG 738  
DB 80 CACGTCGAGGTCCTCCGCGGCGGCTCCGTCGAGCTTGGCGGTCGCGCCGACGATCAGG 21  
QY 739 ACCATGAAGAGGCTTCGTGC 758  
DB 20 ACCATGAAGAGGCTTCGTGC 1

RESULT 2  
US-09-925-065A-769508/C  
/ Sequence 769508, Application US/09925065A  
/ Publication No. US20040181048A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Wang, David G.  
/ TITLE OF INVENTION: Identification and Mapping of Single  
/ FILE REFERENCE: 108827,135  
/ CURRENT APPLICATION NUMBER: US/09/925,065A  
/ PRIOR FILING DATE: 2001-08-08  
/ PRIOR APPLICATION NUMBER: US 60/243,096  
/ PRIOR FILING DATE: 2000-10-24  
/ PRIOR APPLICATION NUMBER: US 60/252,147  
/ PRIOR FILING DATE: 2000-11-20  
/ PRIOR APPLICATION NUMBER: US 60/250,092  
/ PRIOR FILING DATE: 2000-11-30  
/ PRIOR APPLICATION NUMBER: US 60/261,766  
/ PRIOR FILING DATE: 2001-01-16

/ PRIOR APPLICATION NUMBER: US 60/289,846  
/ PRIOR FILING DATE: 2001-05-09  
/ NUMBER OF SEQ ID NOS: 957086  
/ SOFTWARE: FastSeq for Windows Version 4.0  
/ SEQ ID NO 769508  
/ LENGTH: 557  
/ TYPE: DNA  
/ ORGANISM: Homo sapiens  
US-09-925-065A-769508

Query Match 16.2%; Score 506; DB 7; Length 557;  
Best Local Similarity 99.8%; Pred. No. 2,6e-69;  
Matches 556; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 202 CACAGGCAAAATCTCCGCGGAGCTGCGCCCGCTTTTACCTGAGGCTCAAGTTTCCC 261  
DB 557 CACAGGCAAAATCTCCGCGGAGCTGCGCCCGCTTTTACCTGAGGCTCAAGTTTCCC 498  
QY 262 CCGTAAATAGAACTGGGTTGATCTCCGAGCGCTAACTTCAGAACTCGATGGG 321  
DB 497 CCGTAAATAGAACTGGGTTGATCTCCGAGCGCTAACTTCAGAACTCGATGGG 438  
QY 322 AAGGAGGAGAGGATGGGCTCACTCAAGTGACTTCCCGCGTGAAGCCCGCTTAA 381  
DB 437 AAGGAGGAGAGGATGGGCTCACTCAAGTGACTTCCCGCGTGAAGCCCGCTTAA 378  
QY 382 CTGATCCAGGGGGTGGCAGCTCCGCGGAGACGAGGGGGTGGGGGCTTAAAGAAAC 441  
DB 377 CTGATCCAGGGGGTGGCAGCTCCGCGGAGACGAGGGGGTGGGGGCTTAAAGAAAC 318  
QY 442 CTACCCGCGCCCTTGGCAGCGCTTAAAGCGGAGCGCGGCTCTGACGCTTCTT 501  
DB 317 CTACCCGCGCCCTTGGCAGCGCTTAAAGCGGAGCGCGGCTCTGACGCTTCTT 258  
QY 502 CCGGAGTTGGCACTCCGAGAGATGGGAGCCGACCTTCAAGTTCCAGAGAGCCACCT 561  
DB 257 CCGGAGTTGGCACTCCGAGAGATGGGAGCCGACCTTCAAGTTCCAGAGAGCCACCT 198  
QY 562 GAGAGGCGAGGCGGTCAGAGACAGACGTTGACTCGAGTGGCGGCTTGGAGAGATGGA 621  
DB 197 GAGAGGCGAGGCGGTCAGAGACAGACGTTGACTCGAGTGGCGGCTTGGAGAGATGGA 138  
QY 622 CGAGGAGCGGAGAGCCCTTAACGAGGCTCCCTTGGCGCGCCCGCTCCGAGAGCGCAC 681  
DB 137 CGAGGAGCGGAGAGCCCTTAACGAGGCTCCCTTGGCGCGCCCGCTCCGAGAGCGCAC 78  
QY 682 GTGAGGGTCCCGGCGGCTCTCGTGAAGCTTGGCGGTAGCGCCGAGAGATCAGGACC 741  
DB 77 GTGAGGGTCCCGGCGGCTCTCGTGAAGCTTGGCGGTAGCGCCGAGAGATCAGGACC 18  
QY 742 ATGAAGAGCGTTCTGTGC 758  
DB 17 ATGAAGAGCGTTCTGTGC 1

RESULT 3  
US-09-925-065A-736351/C  
/ Sequence 736351, Application US/09925065A  
/ Publication No. US20040181048A1  
/ GENERAL INFORMATION:  
/ APPLICANT: Wang, David G.  
/ TITLE OF INVENTION: Identification and Mapping of Single  
/ FILE REFERENCE: 108827,135  
/ CURRENT APPLICATION NUMBER: US/09/925,065A  
/ PRIOR FILING DATE: 2001-08-08  
/ PRIOR APPLICATION NUMBER: US 60/243,096  
/ PRIOR FILING DATE: 2000-10-24  
/ PRIOR APPLICATION NUMBER: US 60/252,147  
/ PRIOR FILING DATE: 2000-11-20  
/ PRIOR APPLICATION NUMBER: US 60/250,092  
/ PRIOR FILING DATE: 2000-11-30  
/ PRIOR APPLICATION NUMBER: US 60/261,766

;; PRIOR FILING DATE: 2001-01-16  
;; PRIOR APPLICATION NUMBER: US 60/289,846  
;; PRIOR FILING DATE: 2001-05-09  
;; NUMBER OF SEQ ID NOS: 957086  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 736351  
;; LENGTH: 556  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-925-065A-736351

Query Match 16.2%; Score 505; DB 7; Length 556;  
Best Local Similarity 99.8%; Pred. No. 3.7e-69;  
Matches 555; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 203 ACAAGCAAACTCCGCGGAGCTGCGCCCTTTTACCTGAGCCTCAAGTTTCCCATC 262
DB 556 ACAAGCAAACTCCGCGGAGCTGCGCCCTTTTACCTGAGCCTCAAGTTTCCCATC 497
QY 263 CGTAAATATGAACGGGTTGATCTCCGAGCGCTTAACATTCCAGAACTGGATGGGGCGA 322
DB 496 CGTAAATATGAACGGGTTGATCTCCGAGCGCTTAACATTCCAGAACTGGATGGGGCGA 437
QY 323 AGGAGGAGGAGGATGAGCCACCAACCTGACCTCCCGCGTGAAGCCCGCTACAC 382
DB 436 AGGAGGAGGAGGATGAGCCACCAACCTGACCTCCCGCGTGAAGCCCGCTACAC 377
QY 383 TGATCCAGGGGTGCGAGCTCCGCGCGGACGACGCGGGGTGGGCGGCTCTAGAAACCC 442
DB 376 TGATCCAGGGGTGCGAGCTCCGCGCGGACGACGCGGGGTGGGCGGCTCTAGAAACCC 317
QY 443 TACCGCGCGCCCTTGGGAGCGCTTAAGGCGGAGCGCGGCTCTGCAACCTGCTTGGCC 502
DB 316 TACCGCGCGCCCTTGGGAGCGCTTAAGGCGGAGCGCGGCTCTGCAACCTGCTTGGCC 257
QY 503 CGAGTTGGCACCAACGAGAGATGGGAGCCGACCTCAGCTTCGAGGAGACCAACCGT 562
DB 256 CGAGTTGGCACCAACGAGAGATGGGAGCCGACCTCAGCTTCGAGGAGACCAACCGT 197
QY 563 GAGGCGAGGCGGTGACAGACGACGCTGTGACCTCGAGTGGCGCTGGGGAGATGAC 622
DB 196 GAGGCGAGGCGGTGACAGACGACGCTGTGACCTCGAGTGGCGCTGGGGAGATGAC 137
QY 623 GAGGAGGCGGGGAGCCGCTAACGCGGCTCCCTTGGCGGCGCCGCTCCGAGAGGCGAC 682
DB 136 GAGGAGGCGGGGAGCCGCTAACGCGGCTCCCTTGGCGGCGCCGCTCCGAGAGGCGAC 77
QY 683 TCGAGGGTCCCGGCGGAGCTCCGTGACGTTGCGGTAGCGCGCGAGAGTCAACGACCA 742
DB 76 TCGAGGGTCCCGGCGGAGCTCCGTGACGTTGCGGTAGCGCGCGAGAGTCAACGACCA 17
QY 743 TGAAGAGCGTTCTGTGC 758
DB 16 TGAAGAGCGTTCTGTGC 1
```

RESULT 4  
US-09-925-065A-737120  
;; Sequence 737120, Application US/09925065A  
;; Publication No. US20040181048A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Wang, David G.  
;; TITLE OF INVENTION: Identification and Mapping of Single  
;; Nucleotide Polymorphisms in the Human Genome  
;; FILE REFERENCE: 108827.135  
;; CURRENT APPLICATION NUMBER: US/09/925,065A  
;; PRIOR FILING DATE: 2001-08-08  
;; PRIOR APPLICATION NUMBER: US 60/243,096  
;; PRIOR FILING DATE: 2000-10-24  
;; PRIOR APPLICATION NUMBER: US 60/252,147  
;; PRIOR FILING DATE: 2000-11-20  
;; PRIOR APPLICATION NUMBER: US 60/250,092  
;; PRIOR FILING DATE: 2000-11-30

;; PRIOR APPLICATION NUMBER: US 60/261,766  
;; PRIOR FILING DATE: 2001-01-16  
;; PRIOR APPLICATION NUMBER: US 60/289,846  
;; PRIOR FILING DATE: 2001-05-09  
;; NUMBER OF SEQ ID NOS: 957086  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 737120  
;; LENGTH: 555  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-925-065A-737120

Query Match 16.1%; Score 504; DB 7; Length 555;  
Best Local Similarity 99.8%; Pred. No. 5.3e-69;  
Matches 554; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 204 CAGGCAAACTCCGCGGAGCTGCGCCCTTTTACCTGAGCCTCAAGTTTCCCATC 263
DB 1 CAGGCAAACTCCGCGGAGCTGCGCCCTTTTACCTGAGCCTCAAGTTTCCCATC 60
QY 264 GTAATATGAACGGGTTGATCTCCGAGCGCTTAACATTCCAGAACTGGATGGGGCGA 323
DB 61 GTAATATGAACGGGTTGATCTCCGAGCGCTTAACATTCCAGAACTGGATGGGGCGA 120
QY 324 GGGAGGAGGAGGATGAGCCACCAACCTGACCTCCCGCGTGAAGCCCGCTACCACT 383
DB 121 GGGAGGAGGAGGATGAGCCACCAACCTGACCTCCCGCGTGAAGCCCGCTACCACT 180
QY 384 GATCCAGGGGTGCGAGCTCCGCGCGGACGACGCGGGGTGGGCGGCTCTAGAAACCT 443
DB 181 GATCCAGGGGTGCGAGCTCCGCGCGGACGACGCGGGGTGGGCGGCTCTAGAAACCT 240
QY 444 ACCGCGCGCCCTTGGGAGCGCTTAAGGCGGAGCGCGGCTCTGCAACCTGCTTGGCC 503
DB 241 ACCGCGCGCCCTTGGGAGCGCTTAAGGCGGAGCGCGGCTCTGCAACCTGCTTGGCC 300
QY 504 GAGTTGGCACCAACGAGAGATGGGAGCCGACCTCAGCTTCGAGGAGACCAACCGT 563
DB 301 GAGTTGGCACCAACGAGAGATGGGAGCCGACCTCAGCTTCGAGGAGACCAACCGT 360
QY 564 AGGCGAGGCGGTGACAGACGACGCTGTGACCTCGAGTGGCGCTGGGGAGATGAC 623
DB 361 AGGCGAGGCGGTGACAGACGACGCTGTGACCTCGAGTGGCGCTGGGGAGATGAC 420
QY 624 AGGAGCGGGGAGCCGCTAACGCGGCTCCCTTGGCGGCGCCGCTCCGAGAGGCGAC 683
DB 421 AGGAGCGGGGAGCCGCTAACGCGGCTCCCTTGGCGGCGCCGCTCCGAGAGGCGAC 480
QY 684 CGAGGGTCCCGGCGGAGCTCCGTGACGTTGCGGTAGCGCGCGAGAGTCAACGACCA 743
DB 481 CGAGGGTCCCGGCGGAGCTCCGTGACGTTGCGGTAGCGCGCGAGAGTCAACGACCA 540
QY 744 GAAAGAGCGTTCTGTGC 758
DB 541 GAAAGAGCGTTCTGTGC 555
```

RESULT 5  
US-10-301-480-20199/c  
;; Sequence 20199, Application US/10301480  
;; Publication No. US20060057564A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Wang, David G.  
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
;; FILE REFERENCE: 108827.137  
;; CURRENT APPLICATION NUMBER: US/10/301,480  
;; PRIOR FILING DATE: 2002-11-21  
;; PRIOR APPLICATION NUMBER: US 10/215,598  
;; PRIOR FILING DATE: 2002-08-09  
;; PRIOR APPLICATION NUMBER: US 60/311,695  
;; PRIOR FILING DATE: 2001-08-10  
;; NUMBER OF SEQ ID NOS: 1226818



SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 20199  
LENGTH: 477  
TYPE: DNA  
ORGANISM: Homo sapien  
US-10-301-480-20199

Query Match 1.7%; Score 53; DB 11; Length 477;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2888 TGAGGACAGGTGATCCTGAGGCCAGAGTTGAGACCAAGCTTGCCAAACT 2940  
DB 102 TGAGGACAGGTGATCCTGAGGCCAGAGTTGAGACCAAGCTTGCCAAACT 50

RESULT 6  
US-10-301-480-633608/c  
Sequence 633608, Application US/10301480  
Publication No. US20060057564A1  
GENERAL INFORMATION:

APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
FILE REFERENCE: 108827.137  
CURRENT APPLICATION NUMBER: US/10/301,480

PRIOR FILING DATE: 2002-11-21  
PRIOR APPLICATION NUMBER: US 10/215,598  
PRIOR FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: US 60/311,695  
NUMBER OF SEQ ID NOS: 1226818  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 633608  
LENGTH: 477  
TYPE: DNA  
ORGANISM: Homo sapien  
US-10-301-480-633608

Query Match 1.7%; Score 53; DB 12; Length 477;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2888 TGAGGACAGGTGATCCTGAGGCCAGAGTTGAGACCAAGCTTGCCAAACT 2940  
DB 102 TGAGGACAGGTGATCCTGAGGCCAGAGTTGAGACCAAGCTTGCCAAACT 50

RESULT 7  
US-10-301-480-20198/c  
Sequence 20198, Application US/10301480  
Publication No. US20060057564A1  
GENERAL INFORMATION:

APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
FILE REFERENCE: 108827.137  
CURRENT APPLICATION NUMBER: US/10/301,480

PRIOR FILING DATE: 2002-11-21  
PRIOR APPLICATION NUMBER: US 10/215,598  
PRIOR FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: US 60/311,695  
NUMBER OF SEQ ID NOS: 1226818  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 20198  
LENGTH: 492  
TYPE: DNA  
ORGANISM: Homo sapien  
US-10-301-480-20198

Query Match 1.7%; Score 53; DB 11; Length 492;  
Best Local Similarity 100.0%; Pred. No. 2;

Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2888 TGAGGACAGGTGATCCTGAGGCCAGAGTTGAGACCAAGCTTGCCAAACT 2940  
DB 161 TGAGGACAGGTGATCCTGAGGCCAGAGTTGAGACCAAGCTTGCCAAACT 109

RESULT 8  
US-10-301-480-633607/c  
Sequence 633607, Application US/10301480  
Publication No. US20060057564A1  
GENERAL INFORMATION:

APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
FILE REFERENCE: 108827.137  
CURRENT APPLICATION NUMBER: US/10/301,480  
CURRENT FILING DATE: 2002-11-21  
PRIOR APPLICATION NUMBER: US 10/215,598  
PRIOR FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: US 60/311,695  
NUMBER OF SEQ ID NOS: 1226818  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 633607  
LENGTH: 492  
TYPE: DNA  
ORGANISM: Homo sapien  
US-10-301-480-633607

Query Match 1.7%; Score 53; DB 12; Length 492;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2888 TGAGGACAGGTGATCCTGAGGCCAGAGTTGAGACCAAGCTTGCCAAACT 2940  
DB 161 TGAGGACAGGTGATCCTGAGGCCAGAGTTGAGACCAAGCTTGCCAAACT 109

RESULT 9  
US-09-925-065A-591947/c  
Sequence 591947, Application US/09925065A  
Publication No. US20040181048A1  
GENERAL INFORMATION:

APPLICANT: Wang, David G.  
TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.135  
CURRENT APPLICATION NUMBER: US/09/925,065A

PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: US 60/243,096  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: US 60/252,147  
PRIOR FILING DATE: 2000-11-20  
PRIOR APPLICATION NUMBER: US 60/250,092  
PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: US 60/261,766  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/289,846  
PRIOR FILING DATE: 2001-05-09  
NUMBER OF SEQ ID NOS: 957086  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 591947  
LENGTH: 559  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-925-065A-591947

Query Match 1.7%; Score 52; DB 7; Length 559;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCACCTGCACCTCAGGCTGGGCAACAGACCAAGACTGTCTTC 3122



Db 164 CAGGATGGCACTGCACTCCAGCTGGGCAACAGAGCAACTCTGTCTC 113

## RESULT 10

US-09-925-065A-768118  
; Sequence 768118, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 768118  
; LENGTH: 574  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-768118

Query Match 1.7%; Score 52; DB 7; Length 574;  
Best Local Similarity 100.0%; Pred. No. 2.6;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCAAGTGGATCACTGAGGCCAGAGATTGAGACCAAGCCTGGCCAACT 2940

Db 235 GAGGCAAGTGGATCACTGAGGCCAGAGATTGAGACCAAGCCTGGCCAACT 286

## RESULT 11

US-09-925-065A-59682/C  
; Sequence 59682, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 59682  
; LENGTH: 592  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-59682

Query Match 1.7%; Score 52; DB 7; Length 592;  
Best Local Similarity 100.0%; Pred. No. 2.5;

Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCAAGTGGATCACTGAGGCCAGAGATTGAGACCAAGCCTGGCCAACT 2940

Db 175 GAGGCAAGTGGATCACTGAGGCCAGAGATTGAGACCAAGCCTGGCCAACT 124

## RESULT 12

US-10-301-480-160920/C  
; Sequence 160920, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 160920  
; LENGTH: 592  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-160920

Query Match 1.7%; Score 52; DB 11; Length 592;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCAAGTGGATCACTGAGGCCAGAGATTGAGACCAAGCCTGGCCAACT 2940

Db 175 GAGGCAAGTGGATCACTGAGGCCAGAGATTGAGACCAAGCCTGGCCAACT 124

## RESULT 13

US-10-301-480-774329/C  
; Sequence 774329, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 774329  
; LENGTH: 592  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-774329

Query Match 1.7%; Score 52; DB 12; Length 592;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGCAAGTGGATCACTGAGGCCAGAGATTGAGACCAAGCCTGGCCAACT 2940

Db 175 GAGGCAAGTGGATCACTGAGGCCAGAGATTGAGACCAAGCCTGGCCAACT 124

## RESULT 14

US-10-301-480-454942/C

```
/ Sequence 454942, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ PRIOR FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 454942
/ LENGTH: 614
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-454942

Query Match
Best Local Similarity 1.7%; Score 52; DB 12; Length 614;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
DB 306 CAAGATTGTGCGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 255

RESULT 15
US-10-301-480-1068351/c
/ Sequence 1068351, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ PRIOR FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1068351
/ LENGTH: 614
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-1068351

Query Match
Best Local Similarity 1.7%; Score 52; DB 12; Length 614;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
DB 306 CAAGATTGTGCGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 255

RESULT 16
US-09-925-065A-386361/c
/ Sequence 386361, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
```

```
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 386361
/ LENGTH: 617
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-386361

Query Match
Best Local Similarity 1.7%; Score 52; DB 7; Length 617;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
DB 306 CAAGATTGTGCGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 255

RESULT 17
US-10-301-480-611724/c
/ Sequence 611724, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ PRIOR FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 611724
/ LENGTH: 618
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-611724

Query Match
Best Local Similarity 1.7%; Score 52; DB 12; Length 618;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGAGGTGATGATCACTGAGGCGCAGAGTTGAGACCAAGCTGGCCAAAT 2940
DB 599 GAGGAGGTGATGATCACTGAGGCGCAGAGTTGAGACCAAGCTGGCCAAAT 548

RESULT 18
US-10-301-480-1225133/c
/ Sequence 1225133, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ PRIOR FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
```

;; PRIOR FILING DATE: 2001-08-10  
;; NUMBER OF SEQ ID NOS: 122618  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 1225133  
;; LENGTH: 618  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-301-480-1225133

Query Match 1.7%; Score 52; DB 12; Length 618;  
Best Local Similarity 100.0%; Pred. No. 2.5;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGAGGTGATCACTGAGGCGAGATTGAGACCGCTGGCCAAAT 2940  
DB 599 GAGGAGGTGATCACTGAGGCGAGATTGAGACCGCTGGCCAAAT 548

RESULT 19

US-09-925-065A-708701/c  
;; Sequence 708701, Application US/09925065A  
;; Publication No. US20040181048A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Wang, David G.  
;; TITLE OF INVENTION: Identification and Mapping of Single  
;; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome  
;; FILE REFERENCE: 108827.135  
;; CURRENT APPLICATION NUMBER: US/09/925,065A  
;; CURRENT FILING DATE: 2001-08-08  
;; PRIOR APPLICATION NUMBER: US 60/243,096  
;; PRIOR FILING DATE: 2000-10-24  
;; PRIOR APPLICATION NUMBER: US 60/252,147  
;; PRIOR FILING DATE: 2000-11-20  
;; PRIOR APPLICATION NUMBER: US 60/250,092  
;; PRIOR FILING DATE: 2000-11-30  
;; PRIOR APPLICATION NUMBER: US 60/261,766  
;; PRIOR FILING DATE: 2001-01-16  
;; PRIOR APPLICATION NUMBER: US 60/269,846  
;; PRIOR FILING DATE: 2001-05-09  
;; NUMBER OF SEQ ID NOS: 957086  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 708701  
;; LENGTH: 695  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-925-065A-708701

Query Match 1.7%; Score 52; DB 7; Length 695;  
Best Local Similarity 100.0%; Pred. No. 2.3;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGAGGTGATCACTGAGGCGAGATTGAGACCGCTGGCCAAAT 2940  
DB 490 GAGGAGGTGATCACTGAGGCGAGATTGAGACCGCTGGCCAAAT 439

RESULT 20

US-10-301-480-598972  
;; Sequence 598972, Application US/10301480  
;; Publication No. US20060057564A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Wang, David G.  
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
;; TITLE OF INVENTION: In the Human Genome  
;; FILE REFERENCE: 108827.137  
;; CURRENT APPLICATION NUMBER: US/10/301,480  
;; CURRENT FILING DATE: 2002-11-21  
;; PRIOR APPLICATION NUMBER: US 10/215,598  
;; PRIOR FILING DATE: 2002-08-09  
;; PRIOR APPLICATION NUMBER: US 60/311,695  
;; PRIOR FILING DATE: 2001-08-10  
;; NUMBER OF SEQ ID NOS: 122618  
;; SOFTWARE: FastSeq for Windows Version 4.0

;; SEQ ID NO 598972  
;; LENGTH: 974  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-301-480-598972

Query Match 1.7%; Score 52; DB 12; Length 974;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGAGGTGATCACTGAGGCGAGATTGAGACCGCTGGCCAAAT 2940  
DB 916 GAGGAGGTGATCACTGAGGCGAGATTGAGACCGCTGGCCAAAT 967

RESULT 21

US-10-301-480-1212381  
;; Sequence 1212381, Application US/10301480  
;; Publication No. US20060057564A1  
;; GENERAL INFORMATION:  
;; APPLICANT: Wang, David G.  
;; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
;; TITLE OF INVENTION: In the Human Genome  
;; FILE REFERENCE: 108827.137  
;; CURRENT APPLICATION NUMBER: US/10/301,480  
;; CURRENT FILING DATE: 2002-11-21  
;; PRIOR APPLICATION NUMBER: US 10/215,598  
;; PRIOR FILING DATE: 2002-08-09  
;; PRIOR APPLICATION NUMBER: US 60/311,695  
;; PRIOR FILING DATE: 2001-08-10  
;; NUMBER OF SEQ ID NOS: 122618  
;; SOFTWARE: FastSeq for Windows Version 4.0  
;; SEQ ID NO 1212381  
;; LENGTH: 974  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-301-480-1212381

Query Match 1.7%; Score 52; DB 12; Length 974;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2889 GAGGAGGTGATCACTGAGGCGAGATTGAGACCGCTGGCCAAAT 2940  
DB 916 GAGGAGGTGATCACTGAGGCGAGATTGAGACCGCTGGCCAAAT 967

RESULT 22

US-11-121-086-84  
;; Sequence 84, Application US/11121086  
;; Publication No. US20050266459A1  
;; GENERAL INFORMATION:  
;; APPLICANT: POULSEN, TIM S.  
;; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES  
;; FILE REFERENCE: 09138, 6000-00000  
;; CURRENT APPLICATION NUMBER: US/11/121,086  
;; CURRENT FILING DATE: 2005-05-04  
;; PRIOR APPLICATION NUMBER: US 60/567,570  
;; PRIOR FILING DATE: 2004-05-04  
;; NUMBER OF SEQ ID NOS: 107  
;; SOFTWARE: PatentIn version 3.3  
;; SEQ ID NO 84  
;; LENGTH: 118996  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-11-121-086-84

Query Match 1.7%; Score 52; DB 17; Length 118996;  
Best Local Similarity 100.0%; Pred. No. 0.17;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCACATCGACCTCGGCAACAGACCAAGACTCTCTTC 3122

Db 73872 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 73923

RESULT 23  
US-11-121-086-47  
; Sequence 47, Application US/11121086  
; Publication No. US20050266459A1  
; GENERAL INFORMATION:  
; APPLICANT: POULSEN, TIM S.  
; APPLICANT: NIELSEN, KIRSTEN V.  
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES  
; FILE REFERENCE: 09138.6000-00000  
; CURRENT APPLICATION NUMBER: US/11/121,086  
; PRIOR FILING DATE: 2005-05-04  
; PRIOR APPLICATION NUMBER: 60/567,570  
; PRIOR FILING DATE: 2004-05-04  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: PatentIn version 3.3  
; SEQ ID NO 47  
; LENGTH: 137671  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-121-086-47

Query Match 1.7%; Score 52; DB 17; Length 137671;  
Best Local Similarity 100.0%; Pred. No. 0.15;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3071 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
DB 75583 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 75634

RESULT 24  
US-10-995-561-13396/C  
; Sequence 13396, Application US/10995561  
; Publication No. US20050272054A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; FILE REFERENCE: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
; TITLE OF INVENTION: DETECTION AND USES THEREOF  
; FILE REFERENCE: CL001559  
; CURRENT APPLICATION NUMBER: US/10/995,561  
; CURRENT FILING DATE: 2004-11-24  
; NUMBER OF SEQ ID NOS: 85702  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13396  
; LENGTH: 398287  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (1).. (398287)  
; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-  
US-10-995-561-13396

Query Match 1.7%; Score 52; DB 10; Length 398287;  
Best Local Similarity 100.0%; Pred. No. 0.09;  
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3071 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
DB 281811 CAAGATTGTGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 281760

RESULT 25  
US-09-925-065A-211249/C  
; Sequence 211249, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single  
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 211249  
; LENGTH: 385  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-211249

Query Match 1.6%; Score 51; DB 7; Length 385;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2890 AGGCAAGTGATACCTGAGGCCAGAGTTGCAACCGCTGGCCACAT 2940  
DB 167 AGGCAAGTGATACCTGAGGCCAGAGTTGCAACCGCTGGCCACAT 117

RESULT 26  
US-09-925-065A-141833  
; Sequence 141833, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 141833  
; LENGTH: 561  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-141833

Query Match 1.6%; Score 51; DB 7; Length 561;  
Best Local Similarity 100.0%; Pred. No. 3.7;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2890 AGGCAAGTGATACCTGAGGCCAGAGTTGCAACCGCTGGCCACAT 2940  
DB 402 AGGCAAGTGATACCTGAGGCCAGAGTTGCAACCGCTGGCCACAT 452

RESULT 27  
US-10-301-480-236672  
; Sequence 236672, Application US/10301480

```
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US 10/301,480
; PRIORITY FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 236672
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-236672
```

```
Query Match          1.6%; Score 51; DB 12; Length 561;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2890 AGGCAAGTGTGATCACTGAGGCCAGGAGTTGAGACCACTGGCCAAACAT 2940
DB      402 AGGCAAGTGTGATCACTGAGGCCAGGAGTTGAGACCACTGGCCAAACAT 452
```

```
RESULT 28
US-10-301-480-850081
; Sequence 850081, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US 10/301,480
; PRIORITY FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 850081
; LENGTH: 561
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-850081
```

```
Query Match          1.6%; Score 51; DB 12; Length 561;
Best Local Similarity 100.0%; Pred. No. 3.7;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2890 AGGCAAGTGTGATCACTGAGGCCAGGAGTTGAGACCACTGGCCAAACAT 2940
DB      402 AGGCAAGTGTGATCACTGAGGCCAGGAGTTGAGACCACTGGCCAAACAT 452
```

```
RESULT 29
US-10-301-480-1011
; Sequence 1011, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US 10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
```

```
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1011
; LENGTH: 587
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1011
```

```
Query Match          1.6%; Score 51; DB 11; Length 587;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2890 AGGCAAGTGTGATCACTGAGGCCAGGAGTTGAGACCACTGGCCAAACAT 2940
DB      367 AGGCAAGTGTGATCACTGAGGCCAGGAGTTGAGACCACTGGCCAAACAT 417
```

```
RESULT 30
US-10-301-480-42180
; Sequence 42180, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US 10/301,480
; PRIORITY FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 42180
; LENGTH: 587
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-42180
```

```
Query Match          1.6%; Score 51; DB 11; Length 587;
Best Local Similarity 100.0%; Pred. No. 3.6;
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      2890 AGGCAAGTGTGATCACTGAGGCCAGGAGTTGAGACCACTGGCCAAACAT 2940
DB      367 AGGCAAGTGTGATCACTGAGGCCAGGAGTTGAGACCACTGGCCAAACAT 417
```

```
RESULT 31
US-10-301-480-614420
; Sequence 614420, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US 10/301,480
; PRIORITY FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 614420
; LENGTH: 587
; TYPE: DNA
; ORGANISM: Homo sapien
```

US-10-301-480-614420

Query Match 1.6%; Score 51; DB 12; Length 587;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2890 AGGCAAGTGGATCCTGAGGCGGAGGAGTTCCAGACCACTGGCCAACT 2940  
DB 367 AGGCAAGTGGATCCTGAGGCGGAGGAGTTCCAGACCACTGGCCAACT 417

RESULT 32

US-10-301-480-655589  
; Sequence 655589, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 655589  
; LENGTH: 587  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-655589

Query Match 1.6%; Score 51; DB 12; Length 587;  
Best Local Similarity 100.0%; Pred. No. 3.6;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2890 AGGCAAGTGGATCCTGAGGCGGAGGAGTTCCAGACCACTGGCCAACT 2940  
DB 367 AGGCAAGTGGATCCTGAGGCGGAGGAGTTCCAGACCACTGGCCAACT 417

RESULT 33

US-10-301-480-589593/C  
; Sequence 589593, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 589593  
; LENGTH: 852  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-589593

Query Match 1.6%; Score 51; DB 12; Length 852;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGGCCACTGCACTCCAGCTGGGCAACAGAGCAAGCTGTGCT 3121  
DB 494 CAAGATTGGCCACTGCACTCCAGCTGGGCAACAGAGCAAGCTGTGCT 444

RESULT 34

US-10-301-480-1203002/C  
; Sequence 1203002, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1203002  
; LENGTH: 852  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-1203002

Query Match 1.6%; Score 51; DB 12; Length 852;  
Best Local Similarity 100.0%; Pred. No. 3;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGGCCACTGCACTCCAGCTGGGCAACAGAGCAAGCTGTGCT 3121  
DB 494 CAAGATTGGCCACTGCACTCCAGCTGGGCAACAGAGCAAGCTGTGCT 444

RESULT 35

US-11-121-086-107/C  
; Sequence 107, Application US/11121086  
; Publication No. US20050266459A1  
; GENERAL INFORMATION:  
; APPLICANT: POULSEN, TIM S.  
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES  
; FILE REFERENCE: 09138.6000-0000  
; CURRENT APPLICATION NUMBER: US/11/121,086  
; PRIOR FILING DATE: 2005-05-04  
; PRIOR APPLICATION NUMBER: 60/567,570  
; NUMBER OF SEQ ID NOS: 107  
; SOFTWARE: Patent In version 3.3  
; SEQ ID NO 107  
; LENGTH: 197096  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-11-121-086-107

Query Match 1.6%; Score 51; DB 17; Length 197096;  
Best Local Similarity 100.0%; Pred. No. 0.18;  
Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2890 AGGCAAGTGGATCCTGAGGCGGAGGAGTTCCAGACCACTGGCCAACT 2940  
DB 68932 AGGCAAGTGGATCCTGAGGCGGAGGAGTTCCAGACCACTGGCCAACT 68882

RESULT 36

US-09-925-065A-484883/C  
; Sequence 484883, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135

```

; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 484883
; LENGTH: 571
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-484883

Query Match          1.6%; Score 50; DB 7; Length 571;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGACGAACTGTGTCTC 3122
DB      271 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGACGAACTGTGTCTC 222

RESULT 37
US-09-925-065A-785967/c
; Sequence 785967, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 785967
; LENGTH: 572
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-785967

Query Match          1.6%; Score 50; DB 7; Length 572;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGACGAACTGTGTCTC 3122
DB      347 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGACGAACTGTGTCTC 298

RESULT 38
US-09-925-065A-894777/c
; Sequence 894777, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
```

```

; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 894777
; LENGTH: 625
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-894777

Query Match          1.6%; Score 50; DB 7; Length 625;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGACGAACTGTGTCTC 3122
DB      229 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGACGAACTGTGTCTC 180

RESULT 39
US-09-925-065A-892664/c
; Sequence 892664, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 892664
; LENGTH: 630
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-892664

Query Match          1.6%; Score 50; DB 7; Length 630;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGACGAACTGTGTCTC 3122
DB      234 AGATTGTGCACCTGCACTCCAGCTGGGCAACAGACGAACTGTGTCTC 185

RESULT 40
US-09-925-065A-892405/c
; Sequence 892405, Application US/09925065A
```



```
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 892405
LENGTH: 633
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-892405
```

```
Query Match 1.6%; Score 50; DB 7; Length 633;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3073 AGATTGTGCACCTGCAGCTCGGCAACAGACAGACTCTGTCTC 3122
DB 237 AGATTGTGCACCTGCAGCTCGGCAACAGACAGACTCTGTCTC 188
```

```
RESULT 41
US-09-925-065A-892514/c
Sequence 892514, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 892514
LENGTH: 633
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-892514
```

```
Query Match 1.6%; Score 50; DB 7; Length 633;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3073 AGATTGTGCACCTGCAGCTCGGCAACAGACAGACTCTGTCTC 3122
DB 237 AGATTGTGCACCTGCAGCTCGGCAACAGACAGACTCTGTCTC 188
```

```
RESULT 42
US-09-925-065A-949927
Sequence 949927, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 949927
LENGTH: 633
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-949927
```

```
Query Match 1.6%; Score 50; DB 7; Length 633;
Best Local Similarity 100.0%; Pred. No. 4.9;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3073 AGATTGTGCACCTGCAGCTCGGCAACAGACAGACTCTGTCTC 3122
DB 29 AGATTGTGCACCTGCAGCTCGGCAACAGACAGACTCTGTCTC 78
```

```
RESULT 43
US-09-925-065A-917225
Sequence 917225, Application US/09925065A
Publication No. US20040181048A1
GENERAL INFORMATION:
APPLICANT: Wang, David G.
TITLE OF INVENTION: Identification and Mapping of Single
FILE REFERENCE: 108827.135
CURRENT APPLICATION NUMBER: US/09/925,065A
PRIOR FILING DATE: 2001-08-08
PRIOR APPLICATION NUMBER: US 60/243,096
PRIOR FILING DATE: 2000-10-24
PRIOR APPLICATION NUMBER: US 60/252,147
PRIOR FILING DATE: 2000-11-20
PRIOR APPLICATION NUMBER: US 60/250,092
PRIOR FILING DATE: 2000-11-30
PRIOR APPLICATION NUMBER: US 60/261,766
PRIOR FILING DATE: 2001-01-16
PRIOR APPLICATION NUMBER: US 60/289,846
PRIOR FILING DATE: 2001-05-09
NUMBER OF SEQ ID NOS: 957086
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 917225
LENGTH: 676
TYPE: DNA
ORGANISM: Homo sapiens
US-09-925-065A-917225
```

```
Query Match 1.6%; Score 50; DB 7; Length 676;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3073 AGATTGTGCACCTGCAGCTCGGCAACAGACAGACTCTGTCTC 3122
```

```
Db      448 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACGACTGTCTC 497

RESULT 44
US-09-925-065A-917226
; Sequence 917226, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 917226
; LENGTH: 676
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-917226

Query Match      1.6%; Score 50; DB 7; Length 676;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACGACTGTCTC 3122
Db      448 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACGACTGTCTC 497

RESULT 45
US-09-925-065A-944569
; Sequence 944569, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 944569
; LENGTH: 676
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-944569

Query Match      1.6%; Score 50; DB 7; Length 676;
Best Local Similarity 100.0%; Pred. No. 4.7;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACGACTGTCTC 3122
Db      448 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACGACTGTCTC 497
```

```
QY      3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACGACTGTCTC 3122
Db      448 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACGACTGTCTC 497

RESULT 46
US-09-925-065A-934608
; Sequence 934608, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 934608
; LENGTH: 739
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-934608

Query Match      1.6%; Score 50; DB 7; Length 739;
Best Local Similarity 100.0%; Pred. No. 4.5;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACGACTGTCTC 3122
Db      603 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACGACTGTCTC 652

RESULT 47
US-09-925-065A-941367
; Sequence 941367, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 941367
; LENGTH: 742
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-941367
```

Query Match 1.6%; Score 50; DB 7; Length 742;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACTGCACCTCCAGCTGGGCAACAGACAGACTGTGTCTC 3122  
DB 603 AGATTGTGCACTGCACCTCCAGCTGGGCAACAGACAGACTGTGTCTC 652

RESULT 48  
US-10-301-480-567391  
; Sequence 567391, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 567391  
; LENGTH: 999  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-567391

Query Match 1.6%; Score 50; DB 12; Length 999;  
Best Local Similarity 100.0%; Pred. No. 3.9;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACTGCACCTCCAGCTGGGCAACAGACAGACTGTGTCTC 3122  
DB 107 AGATTGTGCACTGCACCTCCAGCTGGGCAACAGACAGACTGTGTCTC 156

RESULT 49  
US-10-301-480-1180800  
; Sequence 1180800, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1180800  
; LENGTH: 999  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-1180800

Query Match 1.6%; Score 50; DB 12; Length 999;  
Best Local Similarity 100.0%; Pred. No. 3.9;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACTGCACCTCCAGCTGGGCAACAGACAGACTGTGTCTC 3122  
DB 107 AGATTGTGCACTGCACCTCCAGCTGGGCAACAGACAGACTGTGTCTC 156

RESULT 50  
US-10-301-480-100096  
; Sequence 100096, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 100096  
; LENGTH: 1986  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-100096

Query Match 1.6%; Score 50; DB 11; Length 1986;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACTGCACCTCCAGCTGGGCAACAGACAGACTGTGTCTC 3122  
DB 197 AGATTGTGCACTGCACCTCCAGCTGGGCAACAGACAGACTGTGTCTC 246

RESULT 51  
US-10-301-480-713505  
; Sequence 713505, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 713505  
; LENGTH: 1986  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-713505

Query Match 1.6%; Score 50; DB 12; Length 1986;  
Best Local Similarity 100.0%; Pred. No. 2.7;  
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACTGCACCTCCAGCTGGGCAACAGACAGACTGTGTCTC 3122  
DB 197 AGATTGTGCACTGCACCTCCAGCTGGGCAACAGACAGACTGTGTCTC 246

RESULT 52  
US-11-121-086-83/C  
; Sequence 83, Application US/11121086  
; Publication No. US20050266459A1  
; GENERAL INFORMATION:  
; APPLICANT: POUlsen, TIM S.  
; APPLICANT: NIELSEN, KRISTEN V.  
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES  
; FILE REFERENCE: 09138.6000-00000

```

; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 83
; LENGTH: 187745
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-83

Query Match          1.6%; Score 50; DB 17; Length 187745;
Best Local Similarity 100.0%; Pred. No. 0.27;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCCTGACCTCCAGCCTGGGCAACAGACCAAGACTGTCTC 3122
DB 159390 AGATTGTGCCTGACCTCCAGCCTGGGCAACAGACCAAGACTGTCTC 159341

RESULT 53
US-11-112-908-21
; Sequence 21, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 21
; LENGTH: 207908
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-21

Query Match          1.6%; Score 50; DB 17; Length 207908;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2888 TGAGGCAAGTGATCACTCCAGGCTGGGCAACAGACCAAGCTGTCTC 2937
DB 168125 TGAGGCAAGTGATCACTCCAGGCTGGGCAACAGACCAAGCTGTCTC 168174

RESULT 54
US-11-112-908-19
; Sequence 19, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harris, Cole
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; CURRENT FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
```

```

; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 19
; LENGTH: 212805
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-19

Query Match          1.6%; Score 50; DB 17; Length 212805;
Best Local Similarity 100.0%; Pred. No. 0.25;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2888 TGAGGCAAGTGATCACTCCAGGCTGGGCAACAGACCAAGCTGTCTC 2937
DB 24046 TGAGGCAAGTGATCACTCCAGGCTGGGCAACAGACCAAGCTGTCTC 24095

RESULT 55
US-10-301-480-235645/C
; Sequence 235645, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 235645
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-235645

Query Match          1.6%; Score 49; DB 12; Length 417;
Best Local Similarity 100.0%; Pred. No. 8.6;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3074 GATTGTGCCACTGCACTCCAGGCTGGGCAACAGACCAAGACTGTCTC 3122
DB 273 GATTGTGCCACTGCACTCCAGGCTGGGCAACAGACCAAGACTGTCTC 225

RESULT 56
US-10-301-480-849054/C
; Sequence 849054, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 849054
; LENGTH: 417
; TYPE: DNA
; ORGANISM: Homo sapiens
```

US-10-301-480-849054

Query Match 1.6%; Score 49; DB 12; Length 417;

Best Local Similarity 100.0%; Pred. No. 8.6; Indels 0; Gaps 0;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3074 GATTGGCAGCTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122

DB 273 GATTGGCAGCTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 225

RESULT 57

US-09-925-065A-140670/c

; Sequence 140670, Application US/09925065A

; Publication No. US20040181048A1

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single

; FILE REFERENCE: 108827.135 Nucleotide Polymorphisms in the Human Genome

; CURRENT APPLICATION NUMBER: US/09/925,065A

; PRIOR FILING DATE: 2001-08-08

; PRIOR APPLICATION NUMBER: US 60/243,096

; PRIOR FILING DATE: 2000-10-24

; PRIOR APPLICATION NUMBER: US 60/252,147

; PRIOR FILING DATE: 2000-11-20

; PRIOR APPLICATION NUMBER: US 60/250,092

; PRIOR FILING DATE: 2000-11-30

; PRIOR APPLICATION NUMBER: US 60/261,766

; PRIOR FILING DATE: 2001-01-16

; PRIOR APPLICATION NUMBER: US 60/289,846

; PRIOR FILING DATE: 2001-05-09

; NUMBER OF SEQ ID NOS: 957086

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 140670

; LENGTH: 420

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-925-065A-140670

Query Match 1.6%; Score 49; DB 7; Length 420;

Best Local Similarity 100.0%; Pred. No. 8.6; Indels 0; Gaps 0;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3074 GATTGGCAGCTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122

DB 273 GATTGGCAGCTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 225

RESULT 58

US-10-301-480-15264/c

; Sequence 15264, Application US/10301480

; Publication No. US20060057564A1

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms

; FILE REFERENCE: 108827.137 in the Human Genome

; CURRENT APPLICATION NUMBER: US/10/301,480

; CURRENT FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US 10/215,598

; PRIOR FILING DATE: 2002-06-09

; PRIOR APPLICATION NUMBER: US 60/311,695

; PRIOR FILING DATE: 2001-08-10

; NUMBER OF SEQ ID NOS: 1226818

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 15264

; LENGTH: 530

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-301-480-15264

Query Match 1.6%; Score 49; DB 11; Length 530;

Best Local Similarity 100.0%; Pred. No. 7.6; Indels 0; Gaps 0;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3074 GATTGGCAGCTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122

DB 230 GATTGGCAGCTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 182

RESULT 59

US-10-301-480-235646/c

; Sequence 235646, Application US/10301480

; Publication No. US20060057564A1

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms

; FILE REFERENCE: 108827.137 in the Human Genome

; CURRENT APPLICATION NUMBER: US/10/301,480

; CURRENT FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US 10/215,598

; PRIOR FILING DATE: 2002-06-09

; PRIOR APPLICATION NUMBER: US 60/311,695

; PRIOR FILING DATE: 2001-08-10

; NUMBER OF SEQ ID NOS: 1226818

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 235646

; LENGTH: 530

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-301-480-235646

Query Match 1.6%; Score 49; DB 12; Length 530;

Best Local Similarity 100.0%; Pred. No. 7.6; Indels 0; Gaps 0;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3074 GATTGGCAGCTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122

DB 230 GATTGGCAGCTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 182

RESULT 60

US-10-301-480-628673/c

; Sequence 628673, Application US/10301480

; Publication No. US20060057564A1

; GENERAL INFORMATION:

; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms

; FILE REFERENCE: 108827.137 in the Human Genome

; CURRENT APPLICATION NUMBER: US/10/301,480

; CURRENT FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US 10/215,598

; PRIOR FILING DATE: 2002-06-09

; PRIOR APPLICATION NUMBER: US 60/311,695

; PRIOR FILING DATE: 2001-08-10

; NUMBER OF SEQ ID NOS: 1226818

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 628673

; LENGTH: 530

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-301-480-628673

Query Match 1.6%; Score 49; DB 12; Length 530;

Best Local Similarity 100.0%; Pred. No. 7.6; Indels 0; Gaps 0;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3074 GATTGGCAGCTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122

DB 230 GATTGGCAGCTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 182

RESULT 61

```
US-10-301-480-849055/c
; Sequence 849055, Application US/10301480
; Publication No. US20060057364A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 122618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 849055
; LENGTH: 530
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-849055
```

```
Query Match 1.6%; Score 49; DB 12; Length 530;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3074 GATTGTGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122
DB 230 GATTGTGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 182
```

```
RESULT 62
US-09-925-065A-140671/c
; Sequence 140671, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 140671
; LENGTH: 532
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-140671
```

```
Query Match 1.6%; Score 49; DB 7; Length 532;
Best Local Similarity 100.0%; Pred. No. 7.6;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3074 GATTGTGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122
DB 232 GATTGTGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 184
```

```
RESULT 63
US-09-925-065A-18609/c
; Sequence 18609, Application US/09925065A
; Publication No. US20040181048A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18609
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-18609
```

```
Query Match 1.6%; Score 49; DB 7; Length 552;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3073 AGATTGTGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCT 3121
DB 363 AGATTGTGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCT 315
```

```
RESULT 64
US-09-925-065A-18610/c
; Sequence 18610, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18610
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-18610
```

```
Query Match 1.6%; Score 49; DB 7; Length 552;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3073 AGATTGTGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCT 3121
DB 363 AGATTGTGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTCTGTCT 315
```

```
RESULT 65
```

```
US-09-925-065A-18611/c
; Sequence 18611, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925, 065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 18611
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-18611
```

```
Query Match      1.6%; Score 49; DB 11; Length 552;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACAAGACTCTGTCT 3121
          |||||||
DB      363 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACAAGACTCTGTCT 315
```

```
RESULT 66
US-10-301-480-119846/c
; Sequence 119846, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 119846
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-119846
```

```
Query Match      1.6%; Score 49; DB 11; Length 552;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACAAGACTCTGTCT 3121
          |||||||
DB      363 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACAAGACTCTGTCT 315
```

```
RESULT 67
US-10-301-480-119847/c
; Sequence 119847, Application US/10301480
; Publication No. US20060057564A1
```

```
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 119847
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-119847
```

```
Query Match      1.6%; Score 49; DB 11; Length 552;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACAAGACTCTGTCT 3121
          |||||||
DB      363 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACAAGACTCTGTCT 315
```

```
RESULT 68
US-10-301-480-119848/c
; Sequence 119848, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 119848
; LENGTH: 552
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-119848
```

```
Query Match      1.6%; Score 49; DB 11; Length 552;
Best Local Similarity 100.0%; Pred. No. 7.5;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3073 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACAAGACTCTGTCT 3121
          |||||||
DB      363 AGATTGTGCACCTGCACCTCCAGCTGGGCAACAGACAAGACTCTGTCT 315
```

```
RESULT 69
US-10-301-480-733255/c
; Sequence 733255, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
```



```
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 733255
/ LENGTH: 552
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-733255

Query Match
Best Local Similarity 100.0%; Pred. No. 7.5; Length 552;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGTGCCACTGCACCTCCAGCCTGGGCAACAGACCAAGACTCTGTCT 3121
Db 363 AGATTGTGCCACTGCACCTCCAGCCTGGGCAACAGACCAAGACTCTGTCT 315

RESULT 70
US-10-301-480-733256/c
/ Sequence 733256, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ PRIOR FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 733256
/ LENGTH: 552
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-733256

Query Match
Best Local Similarity 100.0%; Pred. No. 7.5; Length 552;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGTGCCACTGCACCTCCAGCCTGGGCAACAGACCAAGACTCTGTCT 3121
Db 363 AGATTGTGCCACTGCACCTCCAGCCTGGGCAACAGACCAAGACTCTGTCT 315

RESULT 71
US-10-301-480-733257/c
/ Sequence 733257, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ PRIOR FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 733257
/ LENGTH: 552
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-733257
```

```
Query Match
Best Local Similarity 100.0%; Pred. No. 7.5; Length 552;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGTGCCACTGCACCTCCAGCCTGGGCAACAGACCAAGACTCTGTCT 3121
Db 363 AGATTGTGCCACTGCACCTCCAGCCTGGGCAACAGACCAAGACTCTGTCT 315

RESULT 72
US-09-925-065A-708700/c
/ Sequence 708700, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 708700
/ LENGTH: 695
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-708700

Query Match
Best Local Similarity 100.0%; Pred. No. 6.6; Length 695;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2892 GGAGGTGATGATCCTGAGGCGCAGAGTTGAGACCAAGCCTGGCAACAT 2940
Db 487 GGAGGTGATGATCCTGAGGCGCAGAGTTGAGACCAAGCCTGGCAACAT 439

RESULT 73
US-09-925-065A-932035/c
/ Sequence 932035, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 932035
/ LENGTH: 750
```

```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-932035

Query Match
Best Local Similarity 1.6%; Score 49; DB 7; Length 750;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3074 GATTGGCCACTGCCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122
DB 612 GATTGGCCACTGCCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 564

RESULT 74
US-09-925-065A-932036/C
; Sequence 932036, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 932036
; LENGTH: 750
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-932036

Query Match
Best Local Similarity 1.6%; Score 49; DB 7; Length 750;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3074 GATTGGCCACTGCCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122
DB 612 GATTGGCCACTGCCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 564

RESULT 75
US-10-508-307-14/C
; Sequence 14, Application US/10508307
; Publication No. US20060084790A1
; GENERAL INFORMATION:
; APPLICANT: Stephens, Len
; APPLICANT: Hawkins, Phillip Thomas
; TITLE OF INVENTION: P-REX1, A PTDS (3,4,5) P3-G-BETA-GAMMA-
; TITLE OF INVENTION: REGULATED GUANINE-NUCLEOTIDE EXCHANGE FACTOR
; TITLE OF INVENTION: FOR RAC
; FILE REFERENCE: 130217.401USPC
; CURRENT APPLICATION NUMBER: US/10/508,307
; CURRENT FILING DATE: 2004-09-21
; PRIOR APPLICATION NUMBER: PCT/GB03/01238
; PRIOR FILING DATE: 2003-03-21
; PRIOR APPLICATION NUMBER: GB 0206684.3
; PRIOR FILING DATE: 2002-03-21
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 14
; LENGTH: 243428
; TYPE: DNA
```

```

; ORGANISM: Homo sapiens
US-10-508-307-14

Query Match
Best Local Similarity 1.6%; Score 49; DB 13; Length 243428;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3074 GATTGGCCACTGCCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122
DB 235339 GATTGGCCACTGCCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 235291

RESULT 76
US-10-301-480-24305
; Sequence 24305, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24305
; LENGTH: 437
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-24305

Query Match
Best Local Similarity 1.5%; Score 48; DB 11; Length 437;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3075 ATTGTGCCACTGCCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122
DB 161 ATTGTGCCACTGCCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 208

RESULT 77
US-10-301-480-637714
; Sequence 637714, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 637714
; LENGTH: 437
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-637714

Query Match
Best Local Similarity 1.5%; Score 48; DB 12; Length 437;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3075 ATTGTGCCACTGCCTCCAGCTGGGCAACAGAGCAAGACTCTGTCTC 3122
```

Db 161 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTGTGTCTC 208

## RESULT 78

US-10-301-480-43148  
; Sequence 43148, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 43148  
; LENGTH: 438  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-43148

Query Match 1.5%; Score 48; DB 11; Length 438;  
Best Local Similarity 100.0%; Pred. No. 12;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3075 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTGTGTCTC 3122

Db 161 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTGTGTCTC 208

## RESULT 79

US-10-301-480-656557  
; Sequence 656557, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 656557  
; LENGTH: 438  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-656557

Query Match 1.5%; Score 48; DB 12; Length 438;  
Best Local Similarity 100.0%; Pred. No. 12;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3075 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTGTGTCTC 3122

Db 161 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTGTGTCTC 208

## RESULT 80

US-10-301-480-597620/c  
; Sequence 597620, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms

; FILE REFERENCE: 108827.137

; CURRENT APPLICATION NUMBER: US/10/301,480

; PRIOR FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US 10/215,598

; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: US 60/311,695

; NUMBER OF SEQ ID NOS: 1226818

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 597620

; LENGTH: 617

; TYPE: DNA

; ORGANISM: Homo sapien

US-10-301-480-597620

Query Match 1.5%; Score 48; DB 12; Length 617;  
Best Local Similarity 100.0%; Pred. No. 10;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3075 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTGTGTCTC 3122

Db 379 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTGTGTCTC 332

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms

; FILE REFERENCE: 108827.137

; CURRENT APPLICATION NUMBER: US/10/301,480

; PRIOR FILING DATE: 2002-11-21

; PRIOR APPLICATION NUMBER: US 10/215,598

; PRIOR FILING DATE: 2002-08-09

; PRIOR APPLICATION NUMBER: US 60/311,695

; NUMBER OF SEQ ID NOS: 1226818

; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 1211029

; LENGTH: 617

; TYPE: DNA

; ORGANISM: Homo sapien

US-10-301-480-1211029

Query Match 1.5%; Score 48; DB 12; Length 617;  
Best Local Similarity 100.0%; Pred. No. 10;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3075 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTGTGTCTC 3122

Db 379 ATTGTGCCACTGCACTCCAGCCTGGGCAACAGACAAGACTGTGTCTC 332

## RESULT 82

US-10-301-480-597621/c  
; Sequence 597621, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10

NUMBER OF SEQ ID NOS: 1226818  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 597621  
LENGTH: 802  
TYPE: DNA  
ORGANISM: Homo sapien  
US-10-301-480-597621

Query Match 1.5%; Score 48; DB 12; Length 802;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3075 ATTGGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
DB 379 ATTGGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 332

RESULT 83  
US-10-301-480-1211030/c  
Sequence 1211030, Application US/10301480  
Publication No. US20060057564A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
FILE REFERENCE: 108827.137  
CURRENT APPLICATION NUMBER: US/10/301,480  
PRIOR FILING DATE: 2002-11-21  
PRIOR APPLICATION NUMBER: US 10/215,598  
PRIOR FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: US 60/311,695  
PRIOR FILING DATE: 2001-08-10  
NUMBER OF SEQ ID NOS: 1226818  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 1211030  
LENGTH: 802  
TYPE: DNA  
ORGANISM: Homo sapien  
US-10-301-480-1211030

Query Match 1.5%; Score 48; DB 12; Length 802;  
Best Local Similarity 100.0%; Pred. No. 8.8;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3075 ATTGGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3122  
DB 379 ATTGGCCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 332

RESULT 84  
US-11-124-368A-2878  
Sequence 2878, Application US/11124368A  
Publication No. US20050287559A1  
GENERAL INFORMATION:  
APPLICANT: Michele Cargili  
APPLICANT: James J. Devlin  
APPLICANT: May Luke  
TITLE OF INVENTION: Genetic Polymorphisms Associated with  
FILE REFERENCE: CL001524  
CURRENT APPLICATION NUMBER: US/11/124,368A  
PRIOR FILING DATE: 2005-05-09  
PRIOR APPLICATION NUMBER: US 60/568,845  
PRIOR FILING DATE: 2004-05-07  
PRIOR APPLICATION NUMBER: US 60/625,936  
PRIOR FILING DATE: 2004-11-09  
NUMBER OF SEQ ID NOS: 21112  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 2878  
LENGTH: 19675  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-11-124-368A-2878

Query Match 1.5%; Score 48; DB 17; Length 19675;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3120  
DB 14998 AGATTGGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 15045

RESULT 85  
US-09-925-065A-440136/c  
Sequence 440136, Application US/09925065A  
Publication No. US20040181048A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single  
FILE REFERENCE: 108827.135  
CURRENT APPLICATION NUMBER: US/09/925,065A  
PRIOR FILING DATE: 2001-08-08  
PRIOR APPLICATION NUMBER: US 60/243,096  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: US 60/252,147  
PRIOR FILING DATE: 2000-11-20  
PRIOR APPLICATION NUMBER: US 60/250,092  
PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: US 60/261,766  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/289,846  
PRIOR FILING DATE: 2001-05-09  
NUMBER OF SEQ ID NOS: 957086  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 440136  
LENGTH: 534  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-925-065A-440136

Query Match 1.5%; Score 47; DB 7; Length 534;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 3117  
DB 161 CAAGATTGGCACTGCACTCCAGCTGGGCAACAGAGCAAGACTGTCTC 115

RESULT 86  
US-09-925-065A-440137/c  
Sequence 440137, Application US/09925065A  
Publication No. US20040181048A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single  
FILE REFERENCE: 108827.135  
CURRENT APPLICATION NUMBER: US/09/925,065A  
PRIOR FILING DATE: 2001-08-08  
PRIOR APPLICATION NUMBER: US 60/243,096  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: US 60/252,147  
PRIOR FILING DATE: 2000-11-20  
PRIOR APPLICATION NUMBER: US 60/250,092  
PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: US 60/261,766  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/289,846  
PRIOR FILING DATE: 2001-05-09  
NUMBER OF SEQ ID NOS: 957086  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 440137  
LENGTH: 534

;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-09-925-065A-440137

Query Match 1.5%; Score 47; DB 7; Length 534;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGGCACTGCACCTCCAGCTGGGCAACAGCAAGACTCT 3117  
DB 161 CAAGATTGGCACTGCACCTCCAGCTGGGCAACAGCAAGACTCT 115

RESULT 87  
US-09-925-065A-623360/C  
; Sequence 623360, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 623360  
; LENGTH: 566  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-623360

Query Match 1.5%; Score 47; DB 7; Length 566;  
Best Local Similarity 100.0%; Pred. No. 15;  
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3076 TTGTGCACTGCACCTCCAGCTGGGCAACAGCAAGACTCTGTCTC 3122  
DB 336 TTGTGCACTGCACCTCCAGCTGGGCAACAGCAAGACTCTGTCTC 290

RESULT 88  
US-10-995-561-64399  
; Sequence 64399, Application US/10995561  
; Publication No. US20050272054A1  
; GENERAL INFORMATION:  
; APPLICANT: CARGILL, Michele et al.  
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
; FILE REFERENCE: CL001559  
; CURRENT APPLICATION NUMBER: US/10/995,561  
; PRIOR FILING DATE: 2004-11-24  
; NUMBER OF SEQ ID NOS: 85702  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 64399  
; LENGTH: 201  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-995-561-64399

Query Match 1.5%; Score 46; DB 10; Length 201;  
Best Local Similarity 100.0%; Pred. No. 36;

Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACCCTGAGGCCGAGGATTCCAGACCAAGCTGGCCCAACT 2940  
DB 44 GGTGATCACCCTGAGGCCGAGGATTCCAGACCAAGCTGGCCCAACT 89

RESULT 89  
US-10-301-480-20200/C  
; Sequence 20200, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 20200  
; LENGTH: 496  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-20200

Query Match 1.5%; Score 46; DB 11; Length 496;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACCCTGAGGCCGAGGATTCCAGACCAAGCTGGCCCAACT 2940  
DB 154 GGTGATCACCCTGAGGCCGAGGATTCCAGACCAAGCTGGCCCAACT 109

RESULT 90  
US-10-301-480-633609/C  
; Sequence 633609, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 633609  
; LENGTH: 496  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-633609

Query Match 1.5%; Score 46; DB 12; Length 496;  
Best Local Similarity 100.0%; Pred. No. 23;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACCCTGAGGCCGAGGATTCCAGACCAAGCTGGCCCAACT 2940  
DB 154 GGTGATCACCCTGAGGCCGAGGATTCCAGACCAAGCTGGCCCAACT 109

RESULT 91  
US-10-301-480-570118/C

```
/ Sequence 570118, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ CURRENT FILING DATE: 2002-11-21
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 570118
/ LENGTH: 514
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-570118

Query Match      1.5%; Score 46; DB 12; Length 514;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2895 GGTGATCACTGAGCCAGGAGCTTGAGACCAAGCCTGGCCAACT 2940
DB      343 GGTGATCACTGAGCCAGGAGCTTGAGACCAAGCCTGGCCAACT 298

RESULT 92
/ Sequence 1183527, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ CURRENT FILING DATE: 2002-11-21
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1183527
/ LENGTH: 514
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-1183527

Query Match      1.5%; Score 46; DB 12; Length 514;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2895 GGTGATCACTGAGCCAGGAGCTTGAGACCAAGCCTGGCCAACT 2940
DB      343 GGTGATCACTGAGCCAGGAGCTTGAGACCAAGCCTGGCCAACT 298

RESULT 93
US-10-301-480-577517/c
/ Sequence 577517, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ CURRENT FILING DATE: 2002-11-21
```

```
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 577517
/ LENGTH: 540
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-577517

Query Match      1.5%; Score 46; DB 12; Length 540;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3071 CAAGATTGCGCACTGCACTCCAGCCTGGGCAAGAGCAAGACTC 3116
DB      356 CAAGATTGCGCACTGCACTCCAGCCTGGGCAAGAGCAAGACTC 311

RESULT 94
US-10-301-480-1190926/c
/ Sequence 1190926, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ CURRENT FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1190926
/ LENGTH: 540
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-1190926

Query Match      1.5%; Score 46; DB 12; Length 540;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3071 CAAGATTGCGCACTGCACTCCAGCCTGGGCAAGAGCAAGACTC 3116
DB      356 CAAGATTGCGCACTGCACTCCAGCCTGGGCAAGAGCAAGACTC 311

RESULT 95
US-09-925-065A-124249
/ Sequence 124249, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
```

```

; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 124249
; LENGTH: 559
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-124249

Query Match          1.5%; Score 46; DB 7; Length 559;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2895 GGTGATCACTGAGGCCAGAGATTGAGACCAAGCTGCGCAACAT 2940
DB      358 GGTGATCACTGAGGCCAGAGATTGAGACCAAGCTGCGCAACAT 403

RESULT 96
US-10-301-480-221987/c
; Sequence 221987, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 221987
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-221987

Query Match          1.5%; Score 46; DB 12; Length 568;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2895 GGTGATCACTGAGGCCAGAGATTGAGACCAAGCTGCGCAACAT 2940
DB      212 GGTGATCACTGAGGCCAGAGATTGAGACCAAGCTGCGCAACAT 167

RESULT 97
US-10-301-480-835396/c
; Sequence 835396, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 835396
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-835396
```

```

Query Match          1.5%; Score 46; DB 12; Length 568;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2895 GGTGATCACTGAGGCCAGAGATTGAGACCAAGCTGCGCAACAT 2940
DB      212 GGTGATCACTGAGGCCAGAGATTGAGACCAAGCTGCGCAACAT 167

RESULT 98
US-09-925-065A-843811/c
; Sequence 843811, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 843811
; LENGTH: 606
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-843811

Query Match          1.5%; Score 46; DB 7; Length 606;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2895 GGTGATCACTGAGGCCAGAGATTGAGACCAAGCTGCGCAACAT 2940
DB      298 GGTGATCACTGAGGCCAGAGATTGAGACCAAGCTGCGCAACAT 253

RESULT 99
US-09-925-065A-801712/c
; Sequence 801712, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 801712
; LENGTH: 609
; TYPE: DNA
```



```
/ ORGANISM: Homo sapiens
US-09-925-065A-801712

Query Match
Best Local Similarity 1.5%; Score 46; DB 7; Length 609;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACTGAGGCCAGAGATTGAGACCAGCCTGGCCAAACAT 2940
DB 299 GGTGATCACTGAGGCCAGAGATTGAGACCAGCCTGGCCAAACAT 254

RESULT 100
US-09-925-065A-48663/c
/ Sequence 48663, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 48663
/ LENGTH: 610
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-48663

Query Match
Best Local Similarity 1.5%; Score 46; DB 7; Length 610;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACTGAGGCCAGAGATTGAGACCAGCCTGGCCAAACAT 2940
DB 540 GGTGATCACTGAGGCCAGAGATTGAGACCAGCCTGGCCAAACAT 495

RESULT 101
US-09-925-065A-65044
/ Sequence 65044, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ TITLE OF INVENTION: Nucleotide Polymorphisms in the Human Genome
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 65044
/ LENGTH: 610
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-65044

Query Match
Best Local Similarity 1.5%; Score 46; DB 7; Length 610;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACTGAGGCCAGAGATTGAGACCAGCCTGGCCAAACAT 2940
DB 240 GGTGATCACTGAGGCCAGAGATTGAGACCAGCCTGGCCAAACAT 285

RESULT 102
US-10-301-480-149901/c
/ Sequence 149901, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ TITLE OF INVENTION: In the Human Genome
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ PRIOR FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 149901
/ LENGTH: 610
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-149901

Query Match
Best Local Similarity 1.5%; Score 46; DB 11; Length 610;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACTGAGGCCAGAGATTGAGACCAGCCTGGCCAAACAT 2940
DB 540 GGTGATCACTGAGGCCAGAGATTGAGACCAGCCTGGCCAAACAT 495

RESULT 103
US-10-301-480-166282
/ Sequence 166282, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ TITLE OF INVENTION: In the Human Genome
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ PRIOR FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 166282
/ LENGTH: 610
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-166282

Query Match
Best Local Similarity 1.5%; Score 46; DB 11; Length 610;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy 2895 GGTGATCACTGAGGCCGAGAGTTGAGACCAAGCTTGGCCAACAT 2940  
Db 240 GGTGATCACTGAGGCCGAGAGTTGAGACCAAGCTTGGCCAACAT 285

RESULT 104  
US-10-301-480-763310/C  
; Sequence 763310, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 763310  
; LENGTH: 610  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-301-480-763310

Query Match 1.5%; Score 46; DB 12; Length 610;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2895 GGTGATCACTGAGGCCGAGAGTTGAGACCAAGCTTGGCCAACAT 2940  
Db 540 GGTGATCACTGAGGCCGAGAGTTGAGACCAAGCTTGGCCAACAT 495

RESULT 105  
US-10-301-480-779691  
; Sequence 779691, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; CURRENT FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 779691  
; LENGTH: 610  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-301-480-779691

Query Match 1.5%; Score 46; DB 12; Length 610;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2895 GGTGATCACTGAGGCCGAGAGTTGAGACCAAGCTTGGCCAACAT 2940  
Db 240 GGTGATCACTGAGGCCGAGAGTTGAGACCAAGCTTGGCCAACAT 285

RESULT 106  
US-09-925-065A-47410/C  
; Sequence 47410, Application US/09925065A

; Publication No. US20040181048A1

; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 47410  
; LENGTH: 614  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-47410

Query Match 1.5%; Score 46; DB 7; Length 614;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGCGCACTGCATCTCCAGCTCGGGCAAGAGCAAGACTCTG 3118  
Db 264 AGATTGCGCACTGCATCTCCAGCTCGGGCAAGAGCAAGACTCTG 219

RESULT 107  
US-09-925-065A-47411/C  
; Sequence 47411, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; CURRENT FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 47411  
; LENGTH: 614  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-47411

Query Match 1.5%; Score 46; DB 7; Length 614;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3073 AGATTGCGCACTGCATCTCCAGCTCGGGCAAGAGCAAGACTCTG 3118  
Db 264 AGATTGCGCACTGCATCTCCAGCTCGGGCAAGAGCAAGACTCTG 219

```
RESULT 108
US-10-301-480-148648/c
; Sequence 148648, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 148648
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-148648

Query Match
Best Local Similarity 1.5%; Score 46; DB 11; Length 614;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCGCACTGCAGCTCCAGCTGGGCAACAGAGCAAGACTCTG 3118
DB 264 AGATTGTGCGCACTGCAGCTCCAGCTGGGCAACAGAGCAAGACTCTG 219

RESULT 109
US-10-301-480-148649/c
; Sequence 148649, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 148649
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-148649

Query Match
Best Local Similarity 1.5%; Score 46; DB 11; Length 614;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCGCACTGCAGCTCCAGCTGGGCAACAGAGCAAGACTCTG 3118
DB 264 AGATTGTGCGCACTGCAGCTCCAGCTGGGCAACAGAGCAAGACTCTG 219

RESULT 110
US-10-301-480-762057/c
; Sequence 762057, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
```

```
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 762057
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-762057

Query Match
Best Local Similarity 1.5%; Score 46; DB 12; Length 614;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCGCACTGCAGCTCCAGCTGGGCAACAGAGCAAGACTCTG 3118
DB 264 AGATTGTGCGCACTGCAGCTCCAGCTGGGCAACAGAGCAAGACTCTG 219

RESULT 111
US-10-301-480-762058/c
; Sequence 762058, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 762058
; LENGTH: 614
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-762058

Query Match
Best Local Similarity 1.5%; Score 46; DB 12; Length 614;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCGCACTGCAGCTCCAGCTGGGCAACAGAGCAAGACTCTG 3118
DB 264 AGATTGTGCGCACTGCAGCTCCAGCTGGGCAACAGAGCAAGACTCTG 219

RESULT 112
US-09-925-065A-930543/c
; Sequence 930543, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
```

```

; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 930543
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-930543

Query Match
Best Local Similarity 1.5%; Score 46; DB 7; Length 615;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2897 TGAATCCTGAGGCCAGGAGTTTCGAGACCAAGCTGGCCCAATAG 2942
Db 290 TGAATCCTGAGGCCAGGAGTTTCGAGACCAAGCTGGCCCAATAG 245

RESULT 113
US-09-925-065A-930544/c
; Sequence 930544, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 930544
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-930544

Query Match
Best Local Similarity 1.5%; Score 46; DB 7; Length 615;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2897 TGAATCCTGAGGCCAGGAGTTTCGAGACCAAGCTGGCCCAATAG 2942
Db 290 TGAATCCTGAGGCCAGGAGTTTCGAGACCAAGCTGGCCCAATAG 245

RESULT 114
US-09-925-065A-870312/c
; Sequence 870312, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
```

```

; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 870312
; LENGTH: 616
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-870312

Query Match
Best Local Similarity 1.5%; Score 46; DB 7; Length 616;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3071 CAAGATTGTGCGACTGCACTCGAGCCCTGGGCAAGAGCAAGACTC 3116
Db 246 CAAGATTGTGCGACTGCACTCGAGCCCTGGGCAAGAGCAAGACTC 201

RESULT 115
US-09-925-065A-554324
; Sequence 554324, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 554324
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-554324

Query Match
Best Local Similarity 1.5%; Score 46; DB 7; Length 654;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2895 GGTGATCCTGAGGCCAGGAGTTTCGAGACCAAGCTGGCCCAAT 2940
Db 154 GGTGATCCTGAGGCCAGGAGTTTCGAGACCAAGCTGGCCCAAT 199

RESULT 116
US-09-925-065A-759345/c
; Sequence 759345, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
```

```
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 759345
; LENGTH: 654
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-759345

Query Match      1.5%; Score 46; DB 7; Length 654;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3071 CAAGATTGTGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
DB      274  CAAGATTGTGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 229

RESULT 117
US-09-925-065A-790726
; Sequence 790726, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 790726
; LENGTH: 659
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-790726

Query Match      1.5%; Score 46; DB 7; Length 659;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2895 GGTGATCACCCTGAGGCGAGGAGTTGAGACCAAGCCTGGCCAACAT 2940
DB      180  GGTGATCACCCTGAGGCGAGGAGTTGAGACCAAGCCTGGCCAACAT 225

RESULT 118
US-09-925-065A-790727
; Sequence 790727, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
```

```
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 790727
; LENGTH: 659
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-790727

Query Match      1.5%; Score 46; DB 7; Length 659;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2895 GGTGATCACCCTGAGGCGAGGAGTTGAGACCAAGCCTGGCCAACAT 2940
DB      180  GGTGATCACCCTGAGGCGAGGAGTTGAGACCAAGCCTGGCCAACAT 225

RESULT 119
US-09-925-065A-771454
; Sequence 771454, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 771454
; LENGTH: 661
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-771454

Query Match      1.5%; Score 46; DB 7; Length 661;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3071 CAAGATTGTGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 3116
DB      388  CAAGATTGTGCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTC 433

RESULT 120
US-09-925-065A-737520
; Sequence 737520, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
```

```
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 737520
/ LENGTH: 672
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-737520

Query Match      1.5%; Score 46; DB 7; Length 672;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3071 CAAGATTGTGCCACTGCACTCCAGCTTGGGCAACAGAGAAGACTTC 3116
DB      385 CAAGATTGTGCCACTGCACTCCAGCTTGGGCAACAGAGAAGACTTC 430

RESULT 121
US-09-925-065A-737227/C
/ Sequence 737227, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 737227
/ LENGTH: 680
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-737227

Query Match      1.5%; Score 46; DB 7; Length 680;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 852555
/ LENGTH: 689
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-852555

Query Match      1.5%; Score 46; DB 7; Length 689;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3073 AGATTGTGCCACTGCACTCCAGCTTGGGCAACAGAGAAGACTTCG 3118
DB      293 AGATTGTGCCACTGCACTCCAGCTTGGGCAACAGAGAAGACTTCG 338

RESULT 123
US-09-925-065A-852556
/ Sequence 852556, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 852556
/ LENGTH: 689
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-852556

Query Match      1.5%; Score 46; DB 7; Length 689;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 124  
US-09-925-065A-938958/c  
; Sequence 938958, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 938958  
; LENGTH: 840  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-938958

Query Match 1.5%; Score 46; DB 7; Length 840;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACTCCAGCTTGGGCAACAGACAGACTCTG 3118  
|||||  
DB 146 AGATTGTGCACCTGCACTCCAGCTTGGGCAACAGACAGACTCTG 101

RESULT 125  
US-09-925-065A-938959/c  
; Sequence 938959, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 938959  
; LENGTH: 840  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-938959

Query Match 1.5%; Score 46; DB 7; Length 840;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACTCCAGCTTGGGCAACAGACAGACTCTG 3118  
|||||

DB 146 AGATTGTGCACCTGCACTCCAGCTTGGGCAACAGACAGACTCTG 101

RESULT 126  
US-09-925-065A-938960/c  
; Sequence 938960, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 938960  
; LENGTH: 840  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-938960

Query Match 1.5%; Score 46; DB 7; Length 840;  
Best Local Similarity 100.0%; Pred. No. 17;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACTCCAGCTTGGGCAACAGACAGACTCTG 3118  
|||||  
DB 146 AGATTGTGCACCTGCACTCCAGCTTGGGCAACAGACAGACTCTG 101

RESULT 127  
US-10-301-480-601877/c  
; Sequence 601877, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 122818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 601877  
; LENGTH: 989  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-601877

Query Match 1.5%; Score 46; DB 12; Length 989;  
Best Local Similarity 100.0%; Pred. No. 16;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCACCTGCACTCCAGCTTGGGCAACAGACAGACTCTG 3118  
|||||

DB 852 AGATTGTGCACCTGCACTCCAGCTTGGGCAACAGACAGACTCTG 807



```
RESULT 128
US-10-301-480-1215286/C
; Sequence 1215286, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 122618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1215286
; LENGTH: 989
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1215286

Query Match
Best Local Similarity 100.0%; Score 46; DB 12; Length 989;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AGATTGTGCGACTGCATCTCCAGCTGGCGCAACAGCAAGACTCTG 3118
DB 852 AGATTGTGCGACTGCATCTCCAGCTGGCGCAACAGCAAGACTCTG 807

RESULT 129
US-10-301-480-612422/C
; Sequence 612422, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT FILING DATE: 2002-11-21
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 122618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 612422
; LENGTH: 998
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-612422

Query Match
Best Local Similarity 100.0%; Score 46; DB 12; Length 998;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACCCTGAGGCGAGGAGTTGAGACCAAGCTGGCCAACAT 2940
DB 848 GGTGATCACCCTGAGGCGAGGAGTTGAGACCAAGCTGGCCAACAT 803

RESULT 130
US-10-301-480-1225831/C
; Sequence 1225831, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
```

```
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 122618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1225831
; LENGTH: 998
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1225831

Query Match
Best Local Similarity 100.0%; Score 46; DB 12; Length 998;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACCCTGAGGCGAGGAGTTGAGACCAAGCTGGCCAACAT 2940
DB 848 GGTGATCACCCTGAGGCGAGGAGTTGAGACCAAGCTGGCCAACAT 803

RESULT 131
US-09-925-065A-711495
; Sequence 711495, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT FILING DATE: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 711495
; LENGTH: 1105
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-711495

Query Match
Best Local Similarity 100.0%; Score 46; DB 7; Length 1105;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCGACTGCATCTCCAGCTGGCGCAACAGCAAGACTC 3116
DB 874 CAAGATTGTGCGACTGCATCTCCAGCTGGCGCAACAGCAAGACTC 919

RESULT 132
US-09-925-065A-678054
; Sequence 678054, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT FILING DATE: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
```

```
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 678054
/ LENGTH: 1364
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-678054

Query Match
Best Local Similarity 100.0%; Score 46; DB 7; Length 1364;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3073 AAGATTGTGCACCTGCAGCTTGGGCAACAGAGCAAGACTCTG 3118
DB 351 AAGATTGTGCACCTGCAGCTTGGGCAACAGAGCAAGACTCTG 396

RESULT 133
US-09-925-065A-689300
/ Sequence 689300, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 689300
/ LENGTH: 1635
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-689300

Query Match
Best Local Similarity 100.0%; Score 46; DB 7; Length 1635;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCACCTGCAGCTTGGGCAACAGAGCAAGACTCTG 3116
DB 831 CAAGATTGTGCACCTGCAGCTTGGGCAACAGAGCAAGACTCTG 876

RESULT 134
US-09-925-065A-677066
/ Sequence 677066, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/10/215,598
/ PRIOR FILING DATE: 2002-11-21
```

```
/ CURRENT APPLICATION NUMBER: US/09/925,065A
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 677066
/ LENGTH: 2227
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-677066

Query Match
Best Local Similarity 100.0%; Score 46; DB 7; Length 2227;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACTGAGCCAGAGATTTCGAGACCAAGCTTGGCCAACT 2940
DB 2035 GGTGATCACTGAGCCAGAGATTTCGAGACCAAGCTTGGCCAACT 2080

RESULT 135
US-10-301-480-94903
/ Sequence 94903, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ PRIOR FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1225818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 94903
/ LENGTH: 2798
/ TYPE: DNA
/ ORGANISM: Homo sapien
US-10-301-480-94903

Query Match
Best Local Similarity 100.0%; Score 46; DB 11; Length 2798;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTGCACCTGCAGCTTGGGCAACAGAGCAAGACTCTG 3116
DB 1434 CAAGATTGTGCACCTGCAGCTTGGGCAACAGAGCAAGACTCTG 1479

RESULT 136
US-10-301-480-708312
/ Sequence 708312, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ PRIOR FILING DATE: 2002-11-21
```

;; PRIOR FILING DATE: 2002-08-09  
;; PRIOR APPLICATION NUMBER: US 60/311,695  
;; PRIOR FILING DATE: 2001-08-10  
;; NUMBER OF SEQ ID NOS: 1226818  
;; SOFTWARE: FASTSEQ for Windows Version 4.0  
;; SEQ ID NO 708312  
;; LENGTH: 2798  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
US-10-301-480-708312

Query Match 1.5%; Score 46; DB 12; Length 2798;  
Best Local Similarity 100.0%; Pred. No. 9.3;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTCACCTGACCTGAGCCCTGGCAACAGCAAGACTC 3116  
DB 1434 CAAGATTGTCACCTGACCTGAGCCCTGGCAACAGCAAGACTC 1479

RESULT 137  
US-10-995-561-13369

;; Sequence 13369, Application US/10995561  
;; Publication No. US20050272054A1  
;; GENERAL INFORMATION:

;; APPLICANT: CARGILL, Michele et al.

;; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH

;; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF

;; TITLE OF INVENTION: DETECTION AND USES THEREOF

;; FILE REFERENCE: CL001559

;; CURRENT APPLICATION NUMBER: US/10/995,561

;; CURRENT FILING DATE: 2004-11-24

;; NUMBER OF SEQ ID NOS: 85702

;; SOFTWARE: FASTSEQ for Windows Version 4.0

;; SEQ ID NO 13369

;; LENGTH: 56448

;; TYPE: DNA

;; ORGANISM: Homo sapiens

;; FEATURE:

;; NAME/KEY: misc\_feature

;; LOCATION: (1)...(56448)

;; OTHER INFORMATION: n = A,T,C or G, or insertion/deletion polymorphism (see Tables 1-

US-10-995-561-13369

Query Match 1.5%; Score 46; DB 10; Length 56448;  
Best Local Similarity 100.0%; Pred. No. 2;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 2940  
DB 10994 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 11039

RESULT 138

US-11-112-908-36/c

;; Sequence 36, Application US/11112908

;; Publication No. US20050260659A1

;; GENERAL INFORMATION:

;; APPLICANT: Harris, Cole

;; APPLICANT: Davis, Lisa M.

;; TITLE OF INVENTION: Breast Cancer Biomarkers

;; FILE REFERENCE: 04-164-US

;; CURRENT APPLICATION NUMBER: US/11/112,908

;; CURRENT FILING DATE: 2005-04-22

;; PRIOR APPLICATION NUMBER: US 60/564,758

;; PRIOR FILING DATE: 2004-04-23

;; PRIOR APPLICATION NUMBER: US 60/575,978

;; PRIOR FILING DATE: 2004-06-01

;; PRIOR APPLICATION NUMBER: US 60/631,702

;; PRIOR FILING DATE: 2004-11-30

;; PRIOR APPLICATION NUMBER: US 60/633,826

;; PRIOR FILING DATE: 2004-12-07

;; NUMBER OF SEQ ID NOS: 511

;; SOFTWARE: PatentIn version 3.3  
;; SEQ ID NO 36  
;; LENGTH: 98345  
;; TYPE: DNA  
;; ORGANISM: Homo sapiens  
;; FEATURE:  
;; NAME/KEY: misc\_feature  
;; LOCATION: (49350)..(49350)  
;; OTHER INFORMATION: n is a, c, g, or t  
US-11-112-908-36

Query Match 1.5%; Score 46; DB 17; Length 98345;  
Best Local Similarity 100.0%; Pred. No. 1.5;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGTCACCTGACCTGAGCCCTGGCAACAGCAAGACTC 3116  
DB 18440 CAAGATTGTCACCTGACCTGAGCCCTGGCAACAGCAAGACTC 18395

RESULT 139  
US-10-330-773-795

;; Sequence 795, Application US/10330773

;; Publication No. US20060040262A1

;; GENERAL INFORMATION:

;; APPLICANT: David W. Morris

;; TITLE OF INVENTION: Novel Compositions and Methods in Cancer

;; FILE REFERENCE: 529452001300

;; CURRENT APPLICATION NUMBER: US/10/330,773

;; CURRENT FILING DATE: 2002-12-27

;; NUMBER OF SEQ ID NOS: 981

;; SOFTWARE: FASTSEQ for Windows Version 4.0

;; SEQ ID NO 795

;; LENGTH: 109661

;; TYPE: DNA

;; ORGANISM: Homo sapiens

;; FEATURE:

;; NAME/KEY: misc\_feature

;; LOCATION: (1)...(109661)

;; OTHER INFORMATION: n = A,T,C or G

US-10-330-773-795

Query Match 1.5%; Score 46; DB 11; Length 109661;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2895 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 2940  
DB 40304 GGTGATCACTGAGGCCAGAGTTGAGACCAAGCTGGCCAAACAT 40349

RESULT 140

US-11-112-908-54/c

;; Sequence 54, Application US/11112908

;; Publication No. US20050260659A1

;; GENERAL INFORMATION:

;; APPLICANT: Harris, Cole

;; APPLICANT: Davis, Lisa M.

;; TITLE OF INVENTION: Breast Cancer Biomarkers

;; FILE REFERENCE: 04-164-US

;; CURRENT APPLICATION NUMBER: US/11/112,908

;; CURRENT FILING DATE: 2005-04-22

;; PRIOR APPLICATION NUMBER: US 60/564,758

;; PRIOR FILING DATE: 2004-04-23

;; PRIOR APPLICATION NUMBER: US 60/575,978

;; PRIOR FILING DATE: 2004-06-01

;; PRIOR APPLICATION NUMBER: US 60/631,702

;; PRIOR FILING DATE: 2004-11-30

;; PRIOR APPLICATION NUMBER: US 60/633,826

;; PRIOR FILING DATE: 2004-12-07

;; NUMBER OF SEQ ID NOS: 511

;; SOFTWARE: PatentIn version 3.3

SEQ ID NO 54  
LENGTH: 150450  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-11-112-908-54

Query Match 1.5%; Score 46; DB 17; Length 150450;  
Best Local Similarity 100.0%; Pred. No. 1.2;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGGCCACTGCTCCAGCTGGGCAACAGAGCAAGACTC 3116  
DB 99596 CAAGATTGGCCACTGCTCCAGCTGGGCAACAGAGCAAGACTC 99551

RESULT 141  
US-11-112-908-53/c  
Sequence 53, Application US/11112908  
Publication No. US20050260659A1  
GENERAL INFORMATION:  
APPLICANT: Harris, Cole  
TITLE OF INVENTION: Breast Cancer Biomarkers  
FILE REFERENCE: 04-164-US  
CURRENT APPLICATION NUMBER: US/11/112,908  
CURRENT FILING DATE: 2005-04-22  
PRIOR APPLICATION NUMBER: US 60/564,758  
PRIOR FILING DATE: 2004-04-23  
PRIOR APPLICATION NUMBER: US 60/575,978  
PRIOR FILING DATE: 2004-06-01  
PRIOR APPLICATION NUMBER: US 60/631,702  
PRIOR FILING DATE: 2004-11-30  
PRIOR APPLICATION NUMBER: US 60/633,826  
PRIOR FILING DATE: 2004-12-07  
NUMBER OF SEQ ID NOS: 511  
SOFTWARE: PatentIn version 3.3  
SEQ ID NO 53  
LENGTH: 191343  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-11-112-908-53

Query Match 1.5%; Score 46; DB 17; Length 191343;  
Best Local Similarity 100.0%; Pred. No. 1.1;  
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3071 CAAGATTGGCCACTGCTCCAGCTGGGCAACAGAGCAAGACTC 3116  
DB 18838 CAAGATTGGCCACTGCTCCAGCTGGGCAACAGAGCAAGACTC 18793

RESULT 142  
US-10-995-561-58923/c  
Sequence 58923, Application US/10995561  
Publication No. US20050272054A1  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
FILE REFERENCE: CL001559  
CURRENT APPLICATION NUMBER: US/10/995,561  
CURRENT FILING DATE: 2004-11-24  
NUMBER OF SEQ ID NOS: 85702  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 58923  
LENGTH: 201  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-995-561-58923

Query Match 1.4%; Score 45; DB 10; Length 201;  
Best Local Similarity 100.0%; Pred. No. 51;

Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3078 GTGGCACTGACTCCAGCTGGGCAACAGAGCAAGACTGCTC 3122  
DB 149 GTGGCACTGACTCCAGCTGGGCAACAGAGCAAGACTGCTC 105

RESULT 143  
US-10-995-561-58959/c  
Sequence 58959, Application US/10995561  
Publication No. US20050272054A1  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
FILE REFERENCE: CL001559  
CURRENT APPLICATION NUMBER: US/10/995,561  
CURRENT FILING DATE: 2004-11-24  
NUMBER OF SEQ ID NOS: 85702  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 58959  
LENGTH: 201  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-995-561-58959

Query Match 1.4%; Score 45; DB 10; Length 201;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGGCACTGACTCCAGCTGGGCAACAGAGCAAGACTGCTC 3122  
DB 148 GTGGCACTGACTCCAGCTGGGCAACAGAGCAAGACTGCTC 104

RESULT 144  
US-10-995-561-68537  
Sequence 68537, Application US/10995561  
Publication No. US20050272054A1  
GENERAL INFORMATION:  
APPLICANT: CARGILL, Michele et al.  
TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH  
TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF  
FILE REFERENCE: CL001559  
CURRENT APPLICATION NUMBER: US/10/995,561  
CURRENT FILING DATE: 2004-11-24  
NUMBER OF SEQ ID NOS: 85702  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 68537  
LENGTH: 201  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-995-561-68537

Query Match 1.4%; Score 45; DB 10; Length 201;  
Best Local Similarity 100.0%; Pred. No. 51;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2898 GATCACTGAGGCGGAGAGTTCGAGACCAAGCTGGCCAACTAG 2942  
DB 128 GATCACTGAGGCGGAGAGTTCGAGACCAAGCTGGCCAACTAG 172

RESULT 145  
US-10-301-480-49738/c  
Sequence 49738, Application US/10301480  
Publication No. US20060057564A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
TITLE OF INVENTION: in the Human Genome

```

; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49738
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-49738

Query Match          1.4%; Score 45; DB 11; Length 524;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2888 TGAGGCAAGTGGATCCTGAGGCCGAGAGTTGAGACCAAGCTTG 2932
Db      148 TGAGGCAAGTGGATCCTGAGGCCGAGAGTTGAGACCAAGCTTG 104

RESULT 146
US-10-301-480-49739/c
; Sequence 49739; Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 49739
; LENGTH: 524
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-49739

Query Match          1.4%; Score 45; DB 11; Length 524;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2888 TGAGGCAAGTGGATCCTGAGGCCGAGAGTTGAGACCAAGCTTG 2932
Db      148 TGAGGCAAGTGGATCCTGAGGCCGAGAGTTGAGACCAAGCTTG 104

RESULT 147
US-10-301-480-663147/c
; Sequence 663147; Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 663147
; LENGTH: 524
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-663147

Query Match          1.4%; Score 45; DB 12; Length 524;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2888 TGAGGCAAGTGGATCCTGAGGCCGAGAGTTGAGACCAAGCTTG 2932
Db      148 TGAGGCAAGTGGATCCTGAGGCCGAGAGTTGAGACCAAGCTTG 104

RESULT 148
US-10-301-480-663148/c
; Sequence 663148; Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 663148
; LENGTH: 524
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-663148

Query Match          1.4%; Score 45; DB 12; Length 524;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2888 TGAGGCAAGTGGATCCTGAGGCCGAGAGTTGAGACCAAGCTTG 2932
Db      148 TGAGGCAAGTGGATCCTGAGGCCGAGAGTTGAGACCAAGCTTG 104

RESULT 149
US-10-301-480-9096
; Sequence 9096; Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; CURRENT FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9096
; LENGTH: 527
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-9096

Query Match          1.4%; Score 45; DB 11; Length 527;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

OY 2895 GGTGATCCTGAGGCGAGAGTTGAGACCAAGCCTGGCCACA 2939  
|||||  
DB 55 GGTGATCCTGAGGCGAGAGTTGAGACCAAGCCTGGCCACA 99

RESULT 150  
US-10-301-480-622505  
; Sequence 622505, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 122618  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 622505  
; LENGTH: 527  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-10-301-480-622505

Query Match 1.4%; Score 45; DB 12; Length 527;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2895 GGTGATCCTGAGGCGAGAGTTGAGACCAAGCCTGGCCACA 2939  
|||||  
DB 55 GGTGATCCTGAGGCGAGAGTTGAGACCAAGCCTGGCCACA 99

RESULT 151  
US-09-925-065A-765896  
; Sequence 765896, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 765896  
; LENGTH: 545  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-765896

Query Match 1.4%; Score 45; DB 7; Length 545;  
Best Local Similarity 100.0%; Pred. No. 31;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3078 GTGCCACTGCACTCCAGCTGGGCAACAGACAAGACTGTCTC 3122  
|||||

DB 291 GTGCCACTGCACTCCAGCTGGGCAACAGACAAGACTGTCTC 335

RESULT 152  
US-09-925-065A-799095/C  
; Sequence 799095, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 799095  
; LENGTH: 546  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-799095

Query Match 1.4%; Score 45; DB 7; Length 546;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3078 GTGCCACTGCACTCCAGCTGGGCAACAGACAAGACTGTCTC 3122  
|||||  
DB 210 GTGCCACTGCACTCCAGCTGGGCAACAGACAAGACTGTCTC 166

RESULT 153  
US-09-925-065A-615308  
; Sequence 615308, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 615308  
; LENGTH: 549  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-615308

Query Match 1.4%; Score 45; DB 7; Length 549;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGACGAACTGTCTC 3122  
|||||  
Db 79 GTGCCACTGCACTCCAGCCTGGGCAACAGACGAACTGTCTC 123

## RESULT 154

US-09-925-065A-591948/C  
; Sequence 591948, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 591948  
; LENGTH: 559  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-591948

Query Match 1.4%; Score 45; DB 7; Length 559;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGACGAACTGTCTC 3122  
|||||  
Db 157 GTGCCACTGCACTCCAGCCTGGGCAACAGACGAACTGTCTC 113

## RESULT 155

US-09-925-065A-12438/C  
; Sequence 12438, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single  
; Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 12438  
; LENGTH: 563  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-12438

Query Match 1.4%; Score 45; DB 7; Length 563;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGACGAACTGTCTC 3122  
|||||  
Db 161 GTGCCACTGCACTCCAGCCTGGGCAACAGACGAACTGTCTC 117

## RESULT 156

US-10-301-480-113675/C  
; Sequence 113675, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 113675  
; LENGTH: 563  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-113675

Query Match 1.4%; Score 45; DB 11; Length 563;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGACGAACTGTCTC 3122  
|||||  
Db 161 GTGCCACTGCACTCCAGCCTGGGCAACAGACGAACTGTCTC 117

## RESULT 157

US-10-301-480-727084/C  
; Sequence 727084, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 727084  
; LENGTH: 563  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-727084

Query Match 1.4%; Score 45; DB 12; Length 563;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGACGAACTGTCTC 3122  
|||||  
Db 161 GTGCCACTGCACTCCAGCCTGGGCAACAGACGAACTGTCTC 117



RESULT 158  
US-09-925-065A-13098  
; Sequence 13098, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 13098  
; LENGTH: 568  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-13098

Query Match 1.4%; Score 45; DB 7; Length 568;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTGTCTC 3122  
|||||  
DB 350 GTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTGTCTC 394

RESULT 159  
US-09-925-065A-338452  
; Sequence 338452, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.135  
; CURRENT APPLICATION NUMBER: US/09/925,065A  
; PRIOR FILING DATE: 2001-08-08  
; PRIOR APPLICATION NUMBER: US 60/243,096  
; PRIOR FILING DATE: 2000-10-24  
; PRIOR APPLICATION NUMBER: US 60/252,147  
; PRIOR FILING DATE: 2000-11-20  
; PRIOR APPLICATION NUMBER: US 60/250,092  
; PRIOR FILING DATE: 2000-11-30  
; PRIOR APPLICATION NUMBER: US 60/261,766  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/289,846  
; PRIOR FILING DATE: 2001-05-09  
; NUMBER OF SEQ ID NOS: 957086  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 338452  
; LENGTH: 568  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-925-065A-338452

Query Match 1.4%; Score 45; DB 7; Length 568;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTGTCTC 3122  
|||||

DB 433 GTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTGTCTC 477

RESULT 160  
US-10-301-480-114335  
; Sequence 114335, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 114335  
; LENGTH: 568  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-114335

Query Match 1.4%; Score 45; DB 11; Length 568;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTGTCTC 3122  
|||||  
DB 350 GTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTGTCTC 394

RESULT 161  
US-10-301-480-411598  
; Sequence 411598, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; PRIOR FILING DATE: 2001-08-10  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 411598  
; LENGTH: 568  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-411598

Query Match 1.4%; Score 45; DB 12; Length 568;  
Best Local Similarity 100.0%; Pred. No. 30;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTGTCTC 3122  
|||||  
DB 433 GTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTGTCTC 477

RESULT 162  
US-10-301-480-727744  
; Sequence 727744, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.

```

; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 727744
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-727744

Query Match          1.4%; Score 45; DB 12; Length 568;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
Db      350 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 394

RESULT 163
US-10-301-480-1025007
; Sequence 1025007, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1025007
; LENGTH: 568
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-1025007

Query Match          1.4%; Score 45; DB 12; Length 568;
Best Local Similarity 100.0%; Pred. No. 30;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
Db      433 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 477

RESULT 164
US-09-925-065A-371064/c
; Sequence 371064, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
```

```

; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 371064
; LENGTH: 590
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-371064

Query Match          1.4%; Score 45; DB 7; Length 590;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
Db      433 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 389

RESULT 165
US-09-925-065A-927764
; Sequence 927764, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 927764
; LENGTH: 592
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-927764

Query Match          1.4%; Score 45; DB 7; Length 592;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      3078 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 3122
Db      61 GTGCCACTGCACTCCAGCCTGGGCAACAGAGCAAGACTCTGTCTC 105

RESULT 166
US-09-925-065A-927765
; Sequence 927765, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
```

```

; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 927765
; LENGTH: 592
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-927765

Query Match
1.4%; Score 45; DB 7; Length 592;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3078 GTGCCACTGCATCTCCAGCTGGGCAACAGACAAAGCTGTCTC 3122
|||||
DB 61 GTGCCACTGCATCTCCAGCTGGGCAACAGACAAAGCTGTCTC 105
```

```

RESULT 167
US-09-925-065A-949474
; Sequence 949474, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 949474
; LENGTH: 592
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-949474
```

```

Query Match
1.4%; Score 45; DB 7; Length 592;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3078 GTGCCACTGCATCTCCAGCTGGGCAACAGACAAAGCTGTCTC 3122
|||||
DB 61 GTGCCACTGCATCTCCAGCTGGGCAACAGACAAAGCTGTCTC 105
```

```

RESULT 168
US-09-925-065A-695847/c
; Sequence 695847, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
```

```

; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 695847
; LENGTH: 596
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-695847

Query Match
1.4%; Score 45; DB 7; Length 596;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2906 TGAGGCGAGAGATTGAGACGAGCTGGGCAACATGACAAACC 2950
|||||
DB 508 TGAGGCGAGAGATTGAGACGAGCTGGGCAACATGACAAACC 464
```

```

RESULT 169
US-09-925-065A-10892/c
; Sequence 10892, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; CURRENT FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10892
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-10892
```

```

Query Match
1.4%; Score 45; DB 7; Length 598;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2896 GTGATTCACCTGAGGCGAGAGATTGAGACCGAGCTGGGCAACAT 2940
|||||
DB 304 GTGATTCACCTGAGGCGAGAGATTGAGACCGAGCTGGGCAACAT 260
```

```

RESULT 170
US-10-301-480-112129/c
; Sequence 112129, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
```

```
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 112129
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-112129

Query Match          1.4%; Score 45; DB 11; Length 598;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2896 GTGATCAGCTGAGGCGCAGAGTTGAGACCGCCTGGCCAAACAT 2940
DB      304 GTGATCAGCTGAGGCGCAGAGTTGAGACCGCCTGGCCAAACAT 260

RESULT 171
US-10-301-480-725538/C
; Sequence 725538, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 725538
; LENGTH: 598
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-725538

Query Match          1.4%; Score 45; DB 12; Length 598;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      2896 GTGATCAGCTGAGGCGCAGAGTTGAGACCGCCTGGCCAAACAT 2940
DB      304 GTGATCAGCTGAGGCGCAGAGTTGAGACCGCCTGGCCAAACAT 260

RESULT 172
US-10-301-480-253035/C
; Sequence 253035, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
```

```
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 253035
; LENGTH: 603
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-253035

Query Match          1.4%; Score 45; DB 12; Length 603;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3078 GTGCCACTGCACTCCAGCCTGGGCAAGAGCAAGACTGTCTC 3122
DB      98 GTGCCACTGCACTCCAGCCTGGGCAAGAGCAAGACTGTCTC 54

RESULT 173
US-10-301-480-866444/C
; Sequence 866444, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 866444
; LENGTH: 603
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-866444

Query Match          1.4%; Score 45; DB 12; Length 603;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3078 GTGCCACTGCACTCCAGCCTGGGCAAGAGCAAGACTGTCTC 3122
DB      98 GTGCCACTGCACTCCAGCCTGGGCAAGAGCAAGACTGTCTC 54

RESULT 174
US-09-925-065A-160227/C
; Sequence 160227, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 160227
```

```

; LENGTH: 604
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-160227

```

```

Query Match      1.4%; Score 45; DB 7; Length 604;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3078 GTGCCACTGCATCTCCAGCTGGGCAACAGCAAGACTGTCTC 3122
          |||||||
DB      99 GTGCCACTGCATCTCCAGCTGGGCAACAGCAAGACTGTCTC 55

```

```

RESULT 175
US-09-925-065A-799096/c
; Sequence 799096, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 799096
; LENGTH: 604
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-799096

```

```

Query Match      1.4%; Score 45; DB 7; Length 604;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3078 GTGCCACTGCATCTCCAGCTGGGCAACAGCAAGACTGTCTC 3122
          |||||||
DB      246 GTGCCACTGCATCTCCAGCTGGGCAACAGCAAGACTGTCTC 202

```

```

RESULT 176
US-09-925-065A-795229/c
; Sequence 795229, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09

```

```

; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 795229
; LENGTH: 605
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-795229

```

```

Query Match      1.4%; Score 45; DB 7; Length 605;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3078 GTGCCACTGCATCTCCAGCTGGGCAACAGCAAGACTGTCTC 3122
          |||||||
DB      71 GTGCCACTGCATCTCCAGCTGGGCAACAGCAAGACTGTCTC 27

```

```

RESULT 177

```

```

US-09-925-065A-850785/c
; Sequence 850785, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 850785
; LENGTH: 605
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-850785

```

```

Query Match      1.4%; Score 45; DB 7; Length 605;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      3078 GTGCCACTGCATCTCCAGCTGGGCAACAGCAAGACTGTCTC 3122
          |||||||
DB      71 GTGCCACTGCATCTCCAGCTGGGCAACAGCAAGACTGTCTC 27

```

```

RESULT 178
US-10-301-480-441075/c
; Sequence 441075, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 441075

```

```

; LENGTH: 606
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-441075
```

```
Query Match
Best Local Similarity 100.0%; Pred. No. 29;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGCAAGACTGTCTC 3122
DB 433 GTGCCACTGCACTCCAGCCTGGGCAACAGCAAGACTGTCTC 389
```

```
RESULT 179
US-10-301-480-1054484/c
; Sequence 1054484, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; PRIOR FILING DATE: 2001-08-10
; NUMBER OF SEQ ID NOS: 122618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1054484
; LENGTH: 606
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-301-480-1054484
```

```
Query Match
Best Local Similarity 100.0%; Pred. No. 29;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGCAAGACTGTCTC 3122
DB 433 GTGCCACTGCACTCCAGCCTGGGCAACAGCAAGACTGTCTC 389
```

```
RESULT 180
US-09-925-065A-846491/c
; Sequence 846491, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 846491
; LENGTH: 624
; TYPE: DNA
; ORGANISM: Homo sapiens
```

```
US-09-925-065A-846491
```

```
Query Match
Best Local Similarity 100.0%; Pred. No. 28;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2896 GTGATCACTGAGGCGAGAGTTGAGACCAAGCTGGCCACAT 2940
DB 399 GTGATCACTGAGGCGAGAGTTGAGACCAAGCTGGCCACAT 355
```

```
RESULT 181
US-09-925-065A-839499/c
; Sequence 839499, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 839499
; LENGTH: 627
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-839499
```

```
Query Match
Best Local Similarity 100.0%; Pred. No. 28;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 2896 GTGATCACTGAGGCGAGAGTTGAGACCAAGCTGGCCACAT 2940
DB 402 GTGATCACTGAGGCGAGAGTTGAGACCAAGCTGGCCACAT 358
```

```
RESULT 182
US-09-925-065A-675355/c
; Sequence 675355, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 675355
```

LENGTH: 662  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-925-065A-675355

Query Match 1.4%; Score 45; DB 7; Length 662;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTCTGTCTC 3122  
DB 260 GTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTCTGTCTC 216

RESULT 183  
US-09-925-065A-675356/C  
Sequence 675356, Application US/09925065A  
Publication No. US20040181048A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.135  
CURRENT APPLICATION NUMBER: US/09/925,065A  
PRIOR FILING DATE: 2001-08-08  
PRIOR APPLICATION NUMBER: US 60/243,096  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: US 60/252,147  
PRIOR FILING DATE: 2000-11-20  
PRIOR APPLICATION NUMBER: US 60/250,092  
PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: US 60/261,766  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/289,846  
PRIOR FILING DATE: 2001-05-09  
NUMBER OF SEQ ID NOS: 957086  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 675356  
LENGTH: 662  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-925-065A-675356

Query Match 1.4%; Score 45; DB 7; Length 662;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTCTGTCTC 3122  
DB 260 GTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTCTGTCTC 216

RESULT 184  
US-10-301-480-34229  
Sequence 34229, Application US/10301480  
Publication No. US20060057564A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.137  
CURRENT APPLICATION NUMBER: US/10/301,480  
PRIOR FILING DATE: 2002-11-21  
PRIOR APPLICATION NUMBER: US 10/215,598  
PRIOR FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: US 60/311,695  
PRIOR FILING DATE: 2001-08-10  
NUMBER OF SEQ ID NOS: 1226818  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 34229  
LENGTH: 713  
TYPE: DNA  
ORGANISM: Homo sapiens

US-10-301-480-34229

Query Match 1.4%; Score 45; DB 11; Length 713;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3077 TGTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTCTGTCT 3121  
DB 468 TGTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTCTGTCT 512

RESULT 185  
US-10-301-480-647638  
Sequence 647638, Application US/10301480  
Publication No. US20060057564A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.137  
CURRENT APPLICATION NUMBER: US/10/301,480  
PRIOR FILING DATE: 2002-11-21  
PRIOR APPLICATION NUMBER: US 10/215,598  
PRIOR FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: US 60/311,695  
PRIOR FILING DATE: 2001-08-10  
NUMBER OF SEQ ID NOS: 1226818  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 647638  
LENGTH: 713  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-301-480-647638

Query Match 1.4%; Score 45; DB 12; Length 713;  
Best Local Similarity 100.0%; Pred. No. 27;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3077 TGTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTCTGTCT 3121  
DB 468 TGTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTCTGTCT 512

RESULT 186  
US-10-301-480-593973/C  
Sequence 593973, Application US/10301480  
Publication No. US20060057564A1  
GENERAL INFORMATION:  
APPLICANT: Wang, David G.  
TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome  
FILE REFERENCE: 108827.137  
CURRENT APPLICATION NUMBER: US/10/301,480  
PRIOR FILING DATE: 2002-11-21  
PRIOR APPLICATION NUMBER: US 10/215,598  
PRIOR FILING DATE: 2002-08-09  
PRIOR APPLICATION NUMBER: US 60/311,695  
PRIOR FILING DATE: 2001-08-10  
NUMBER OF SEQ ID NOS: 1226818  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 593973  
LENGTH: 723  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-301-480-593973

Query Match 1.4%; Score 45; DB 12; Length 723;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTCTGTCTC 3122  
DB 113 GTGCCACTGCACTCCAGCTGGGCAACAGCAAGACTCTGTCTC 69



```
RESULT 187
US-10-301-480-1207382/c
; Sequence 1207382, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1207382
; LENGTH: 723
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1207382

Query Match      1.4%; Score 45; DB 12; Length 723;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3078 GTGCCACTGCACTCCAGCTCGGCAACAGAGCAAGACTGTGTCTC 3122
DB      113 GTGCCACTGCACTCCAGCTCGGCAACAGAGCAAGACTGTGTCTC 69

RESULT 188
US-10-301-480-572562
; Sequence 572562, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 572562
; LENGTH: 745
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-572562

Query Match      1.4%; Score 45; DB 12; Length 745;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3078 GTGCCACTGCACTCCAGCTCGGCAACAGAGCAAGACTGTGTCTC 3122
DB      560 GTGCCACTGCACTCCAGCTCGGCAACAGAGCAAGACTGTGTCTC 604

RESULT 189
US-10-301-480-1185971
; Sequence 1185971, Application US/10301480
; Publication No. US20060057564A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
```

```
; TITLE OF INVENTION: In the Human Genome
; FILE REFERENCE: 108827.137
; CURRENT APPLICATION NUMBER: US/10/301,480
; PRIOR FILING DATE: 2002-11-21
; PRIOR APPLICATION NUMBER: US 10/215,598
; PRIOR FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: US 60/311,695
; NUMBER OF SEQ ID NOS: 1226818
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1185971
; LENGTH: 745
; TYPE: DNA
; ORGANISM: Homo sapien
US-10-301-480-1185971

Query Match      1.4%; Score 45; DB 12; Length 745;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3078 GTGCCACTGCACTCCAGCTCGGCAACAGAGCAAGACTGTGTCTC 3122
DB      560 GTGCCACTGCACTCCAGCTCGGCAACAGAGCAAGACTGTGTCTC 604

RESULT 190
US-09-925-065A-13524/C
; Sequence 13524, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
; PRIOR FILING DATE: 2001-08-08
; PRIOR APPLICATION NUMBER: US 60/243,096
; PRIOR FILING DATE: 2000-10-24
; PRIOR APPLICATION NUMBER: US 60/252,147
; PRIOR FILING DATE: 2000-11-20
; PRIOR APPLICATION NUMBER: US 60/250,092
; PRIOR FILING DATE: 2000-11-30
; PRIOR APPLICATION NUMBER: US 60/261,766
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/289,846
; PRIOR FILING DATE: 2001-05-09
; NUMBER OF SEQ ID NOS: 957086
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13524
; LENGTH: 793
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-065A-13524

Query Match      1.4%; Score 45; DB 7; Length 793;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3078 GTGCCACTGCACTCCAGCTCGGCAACAGAGCAAGACTGTGTCTC 3122
DB      740 GTGCCACTGCACTCCAGCTCGGCAACAGAGCAAGACTGTGTCTC 696

RESULT 191
US-09-925-065A-77588/C
; Sequence 77588, Application US/09925065A
; Publication No. US20040181048A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.135
; CURRENT APPLICATION NUMBER: US/09/925,065A
```

```
/ CURRENT FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243,096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252,147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250,092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261,766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/289,846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 77588
/ LENGTH: 793
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-77588
```

```
Query Match      1.4%; Score 45; DB 7; Length 793;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3078 GTGCCACTGCACCTCCAGCCTGGGCAACAGACAAAGACTGTCTC 3122
Db      740 GTGCCACTGCACCTCCAGCCTGGGCAACAGACAAAGACTGTCTC 696
```

```
RESULT 192
US-10-301-480-114761/c
/ Sequence 114761, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ PRIOR FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 114761
/ LENGTH: 793
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-301-480-114761
```

```
Query Match      1.4%; Score 45; DB 11; Length 793;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3078 GTGCCACTGCACCTCCAGCCTGGGCAACAGACAAAGACTGTCTC 3122
Db      740 GTGCCACTGCACCTCCAGCCTGGGCAACAGACAAAGACTGTCTC 696
```

```
RESULT 193
US-10-301-480-178827/c
/ Sequence 178827, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ PRIOR FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
```

```
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 178827
/ LENGTH: 793
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-301-480-178827
```

```
Query Match      1.4%; Score 45; DB 11; Length 793;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3078 GTGCCACTGCACCTCCAGCCTGGGCAACAGACAAAGACTGTCTC 3122
Db      740 GTGCCACTGCACCTCCAGCCTGGGCAACAGACAAAGACTGTCTC 696
```

```
RESULT 194
US-10-301-480-728170/c
/ Sequence 728170, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ PRIOR FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 728170
/ LENGTH: 793
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-301-480-728170
```

```
Query Match      1.4%; Score 45; DB 12; Length 793;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      3078 GTGCCACTGCACCTCCAGCCTGGGCAACAGACAAAGACTGTCTC 3122
Db      740 GTGCCACTGCACCTCCAGCCTGGGCAACAGACAAAGACTGTCTC 696
```

```
RESULT 195
US-10-301-480-792236/c
/ Sequence 792236, Application US/10301480
/ Publication No. US20060057564A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms
/ FILE REFERENCE: 108827.137
/ CURRENT APPLICATION NUMBER: US/10/301,480
/ PRIOR FILING DATE: 2002-11-21
/ PRIOR APPLICATION NUMBER: US 10/215,598
/ PRIOR FILING DATE: 2002-08-09
/ PRIOR APPLICATION NUMBER: US 60/311,695
/ PRIOR FILING DATE: 2001-08-10
/ NUMBER OF SEQ ID NOS: 1226818
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 792236
/ LENGTH: 793
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-10-301-480-792236
```

Query Match 1.4%; Score 45; DB 12; Length 793;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3078 GTGCCACTGCACTCCAGCTTGCGCAACAGAGCAAGACTCTGTCTC 3122  
DB 740 GTGCCACTGCACTCCAGCTTGCGCAACAGAGCAAGACTCTGTCTC 696

RESULT 196  
US-10-301-480-556511/c  
; Sequence 556511, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 556511  
; LENGTH: 803  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-556511

Query Match 1.4%; Score 45; DB 12; Length 803;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3078 GTGCCACTGCACTCCAGCTTGCGCAACAGAGCAAGACTCTGTCTC 3122  
DB 186 GTGCCACTGCACTCCAGCTTGCGCAACAGAGCAAGACTCTGTCTC 142

RESULT 197  
US-10-301-480-1169920/c  
; Sequence 1169920, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1169920  
; LENGTH: 803  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-1169920

Query Match 1.4%; Score 45; DB 12; Length 803;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3078 GTGCCACTGCACTCCAGCTTGCGCAACAGAGCAAGACTCTGTCTC 3122  
DB 186 GTGCCACTGCACTCCAGCTTGCGCAACAGAGCAAGACTCTGTCTC 142

RESULT 198  
US-10-301-480-556318/c  
; Sequence 556318, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 556318  
; LENGTH: 818  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-556318

Query Match 1.4%; Score 45; DB 12; Length 818;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3078 GTGCCACTGCACTCCAGCTTGCGCAACAGAGCAAGACTCTGTCTC 3122  
DB 186 GTGCCACTGCACTCCAGCTTGCGCAACAGAGCAAGACTCTGTCTC 142

RESULT 199  
US-10-301-480-1169727/c  
; Sequence 1169727, Application US/10301480  
; Publication No. US20060057564A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms  
; FILE REFERENCE: 108827.137  
; CURRENT APPLICATION NUMBER: US/10/301,480  
; PRIOR FILING DATE: 2002-11-21  
; PRIOR APPLICATION NUMBER: US 10/215,598  
; PRIOR FILING DATE: 2002-08-09  
; PRIOR APPLICATION NUMBER: US 60/311,695  
; NUMBER OF SEQ ID NOS: 1226818  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1169727  
; LENGTH: 818  
; TYPE: DNA  
; ORGANISM: Homo sapien  
US-10-301-480-1169727

Query Match 1.4%; Score 45; DB 12; Length 818;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3078 GTGCCACTGCACTCCAGCTTGCGCAACAGAGCAAGACTCTGTCTC 3122  
DB 186 GTGCCACTGCACTCCAGCTTGCGCAACAGAGCAAGACTCTGTCTC 142

RESULT 200  
US-09-925-065A-88887  
; Sequence 88887, Application US/09925065A  
; Publication No. US20040181048A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, David G.  
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide Polymorphisms in the Human Genome

FILE REFERENCE: 108827.135  
CURRENT APPLICATION NUMBER: US/09/925,065A  
CURRENT FILING DATE: 2001-08-08  
PRIOR APPLICATION NUMBER: US 60/243,096  
PRIOR FILING DATE: 2000-10-24  
PRIOR APPLICATION NUMBER: US 60/252,147  
PRIOR FILING DATE: 2000-11-20  
PRIOR APPLICATION NUMBER: US 60/250,092  
PRIOR FILING DATE: 2000-11-30  
PRIOR APPLICATION NUMBER: US 60/261,766  
PRIOR FILING DATE: 2001-01-16  
PRIOR APPLICATION NUMBER: US 60/289,846  
PRIOR FILING DATE: 2001-05-09  
NUMBER OF SEQ ID NOS: 957086  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO: 88887  
LENGTH: 896  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-925-065A-88887

Query Match 1.4%; Score 45; DB 7; Length 896;  
Best Local Similarity 100.0%; Pred. No. 24;  
Matches 45; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3078 GTGCCACTGCACTCCAGCCTGGGCAACAGCAAGACTCTGTCTC 3122  
DB 457 GTGCCACTGCACTCCAGCCTGGGCAACAGCAAGACTCTGTCTC 501

Search completed: May 11, 2006, 06:25:33  
Job time : 1350 secs

**THIS PAGE BLANK (USPTO)**